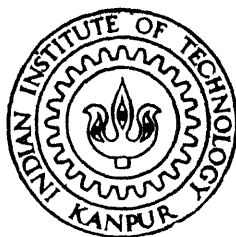


Studies on GPS and its Applications in Networking

by
R. Hemachandra



**EE
1998
M
HEM
STU**

**Department of Electrical Engineering
INDIAN INSTITUTE OF TECHNOLOGY KANPUR
May, 1998**

**Studies on GPS
and
its Applications in Networking**

*A Thesis Submitted
in Partial Fulfillment of the Requirements
for the degree of*

Master of Technology

by
R Hemachandra

to the
**DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY, KANPUR**

May 1998

14 JUL 1998
CENTRAL LIBRARY
IIT KANPUR

Vol. No. A 125737

appl
subm

K
K R
Dep
Indr
(C

Entered in system

EE-1998-M-HEM-STU



A125737



Certificate

This is to certify that the work contained in this thesis entitled *Studies on GPS and its applications in Networking* has been carried out under my supervision and that this work has not been submitted elsewhere for a degree

K R Srivathsan

K R. Srivathsan

Department of Electrical Engineering,

Indian Institute of Technology Kanpur

10-5-98

Acknowledgments

I would like to take this opportunity to express my sincere gratitude to my thesis supervisor Dr K R Srivathsan for his invaluable guidance and constant encouragement throughout the project work. I would also thank him for giving me an opportunity to work in an interesting and new area. I would also use this occasion to express my gratitude to all my teachers with whom I have interacted during the course work.

I would like to thank Amitabha roy for some of his helpful suggestions during the project work. I would also like to thank all the other Research engineers of ERNET Lab - T S Rao, Sandhya, Neeruj, Vivek, Vineet and Kaushal - whose company has made working in this lab such a memorable experience.

Finally I would thank all my friends of Hall-5 for making my stay at IITK thoroughly enjoyable. In particular, special thanks to bhaskar, murty, giridhar, srinivas, kv, maheedhar and sudheer for their excellent company and their readiness to help in times of need.

Abstract

This thesis is a study on some of the ways in which GPS can be put to use in networking. In particular we have employed GPS as an accurate source of UTC time and used this precise GPS timescale to synchronize the clocks of different systems in a computer network. Accurately synchronized clocks are an essential ingredient for the smooth functioning and reliable trouble-shooting of a distributed computer network. The NTP has been widely used to achieve a reliable network time synchronization through a hierarchy of time servers. We have implemented a stratum one NTP Primary time server on a Linux platform using a GPS receiver. This time server can provide an accurate time reference to all the systems within the institute network or outside, if necessary. The thesis also carries a short review of the structure of GPS signals, the commercial grade GPS receiver outputs and their variance in standalone and differential modes.

Contents

1	Introduction	01
1 1	Organization of the Thesis	02
1 2	GPS and Networks	02
2	Overview of the GPS System	05
2 1	Principle of GPS Operation	05
2 2	GPS Signal Characteristics	07
2 3	GPS System Segments	08
2 4	Carrier Modulation	11
2 5	GPS Receiver Structure	14
2 6	Receiver Classification	16
2 7	GPS and Time	19
2 8	GPS Error Sources	20
2 9	Differential GPS	21
2 10	Accords GPS Receiver	22
3	Network Time Synchronization	24
3 1	Computer clocks	26
3 2	The Network Time Protocol	27
3 3	System Architecture	28
3 4	Modes of operation	29
3 5	Data Formats	31

3 6	NTP Functioning Model	32
3 6 1	Delay and Offset Computation	33
3 6 2	Data Filtering	35
3 6 3	Peer Selection and Combining Algorithms	36
3 7	The NTP Local Clock	39
3 8	Present Status	42
4	GPS Time and Position Outputs	
	Studies and some Applications	43
4 1	Implementation of a Time Server	43
4 1 1	Reference Clock Support	43
4 1 2	Running the NTP Daemon	44
4 2	GPS Receiver Measurements A study	48
4 2 1	DOP Measurements	48
4 2 2	Standalone and Differential Position Measurements	53
5	Conclusions	57
5 1	Suggestions for Further work	58
	Appendix A Accords GPS Receiver Outputs	59
	References	62

List of Figures

Fig 2 1	GPS signal structure	12
Fig 2 2	GPS Frame and Subframe format	14
Fig 2 3	GPS Receiver functional Block diagram	15
Fig 2 4	GPSR 2000 Receiver model	23
Fig 3 1	Synchronization subnet	29
Fig 3 2	Modified Synchronization subnet	29
Fig 3 3	NTP Packet Header	32
Fig 3 4	Time server model	33
Fig 3 5	Measuring delay and offset	34
Fig 3 6	NTP clock model	39
Fig 3 7	Update Nomenclature	40
Fig 4 1	Association status	46
Fig 4 2	Clock status	46
Fig 4 3	Section of the messages logged	47
Fig 4 4	PDOP Variation obtained with Receiver 1	49
Fig 4 5	PDOP Variation obtained with Receiver 2	49
Fig 4 6	HDOP Variation obtained with Receiver 1	50
Fig 4 7	HDOP Variation obtained with Receiver 2	50
Fig 4 8	PDOP Variation at Location 1	51
Fig 4 9	HDOP Variation at Location 1	51
Fig 4 10	PDOP Variation at Location 2	52
Fig 4 11	HDOP Variation at Location 2	52
Fig 4 12	GPS Latitude output at Location 1	54
Fig 4 13	GPS Latitude output at Location 2	54
Fig 4 14	Latitude variance at Location 1	55
Fig 4 15	Latitude variance at Location 2	55
Fig 4 16	Differential Latitude variations	56
Fig 4 17	Latitude Variance in differential mode	56

Chapter 1

Introduction

The Global Positioning System (GPS), which has its roots in the cold war era, is turning out to be one of the most exciting and revolutionary developments ever to have taken place. The GPS system was originally designed by the US military in the early 1970s to provide rapid timing and positioning for remote military users such as submarines. However, now the number of civilian uses far outstrip the defence applications. Even though it has been primarily used as a navigation aid in ships and aircraft, new ways to use its capabilities are being continually found. GPS is being used extensively in mapping and surveying. Using GPS to survey and map it precisely saves time and money. In surveying, a line of sight along the ground is no longer necessary for precise positioning. Positions can be measured up to 30 Km from reference points without intermediate points. By combining GPS and computer mapping techniques we will be better able to identify and manage our natural resources. GPS in conjunction with GIS (Geographic Information systems) has opened a whole lot of opportunities. GPS, in combination with digital city maps, is a key component for the Intelligent Vehicle Highway System (IVHS). It has also found applications in as diverse areas as farming, animal tracking, atmospheric experiments and asset management. At the present GPS is the most competent system for time transfer, the distribution of Precise Time and Time Interval (PTTI). This feature of GPS is being utilized by the National timing laboratories. With its capability to disseminate a global time and frequency reference signals, it is being employed in telecom network synchronization and

in improving the efficiency of electric power distribution. Thus the GPS system is set to introduce radical improvements to many systems that are beginning to impact people in all fields.

1 1 Organization of the Thesis

This thesis work has been structured into five chapters. The second chapter presents an overall picture of the GPS system operation and design. Among the things considered are the GPS signal structure, GPS system segments, the functional model of a typical GPS Receiver and the types of receivers. This thesis focuses mainly on the application of GPS for Network synchronization purposes, and thus, this chapter looks in detail at why the GPS serves as a precise time and frequency reference.

The third chapter first considers the role of time synchronization in computer networks. A few applications, where timing is critical, are illustrated. The chapter then proceeds to discuss the Network Time Protocol (NTP), which is currently the most widely used protocol in the Internet community to coordinate the time distribution and to achieve a reliable time synchronization. The NTP architecture, NTP time keeping model and the NTP performance are all considered. The fourth chapter presents the work that we have carried out and is organized into two main sub-sections. The first section considers the incorporation of GPS time in NTP. It presents the implementation details of a GPS based Time server that we have set up. In the next section we present the results of our study on the GPS Receiver outputs. Finally the summary conclusions of the thesis and some suggestions for further work are presented in chapter 5. The rest of this chapter covers some of the ways GPS is being put to use in the networking area.

1 2 GPS and Networks

The satellites that are part of the Global positioning system transmit highly precise time and frequency reference signals. These can be received by relatively inexpensive devices and used as primary reference sources to

synchronize digital communications and timestamp many critical functions within a system. Timing and synchronization are critical for the satisfactory performance of many digital systems. In digital networks, timing problems or slips can seriously degrade the transmission of voice, data or video. In the earlier analog networks, which were tolerant of frequency inaccuracies, cesium clocks were maintained at a few primary centers and these timing signals were then distributed (called the synchronization hierarchy). Newer digital systems however demand, a far more exacting performance. AT&T is using GPS to synchronize its newer generation of digital networks employing the Synchronous optical network (SONET) transmission technology [13]. The GPS timing signal synchronizes a series of master clocks, which in turn produce signals that are distributed throughout the network via a series of primary and secondary nodes. In networks there is a need to accommodate different time delays when multiple bit streams terminate at a node, say for instance a long distance trunk switch. For this purpose, synchronous communications systems relies on accurate frequencies being available.

Wireless service providers who want to get the most out of their limited spectrum are turning to more carefully synchronized systems based on GPS. Current usage of GPS technology include paging systems, CDMA cellular communication systems and mobile platforms such as laptops. Paging systems require synchronized broadcast over their coverage area, otherwise the same page may be received and announced two or three times. In paging systems time synchronization of approximately 10 microseconds (transmitter to transmitter) is typically required, easily provided by the GPS. As the cellular world is going digital (to take advantage of the increased subscriber density, better quality etc.), TDMA and CDMA are the preferred technologies. GPS has provided cost effective solutions for both of them. CDMA which has a superior performance demands stringent synchronization requirements. CDMA cellular communication systems are using GPS for both time synchronization and frequency control. It requires each cell to be within a few microseconds of the CDMA system time base and needs a frequency accuracy of a few parts per billion. To meet these requirements a GPS-disciplined clock is used, which is a

combination of a GPS receiver and a high quality quartz oscillator. The long term stability of GPS complements the short term accuracy of the crystal oscillator (these low cost quartz clocks are very stable but drift away, GPS disciplines this drift).

Wireless services are also benefitting from the GPS broadcast position. They can maintain network maps of their resources, such as broadcast and relay towers, signal strengths etc. The reliability of the service is maintained with mobile test equipment position-tagging the collected data (signal strengths) with GPS. In cellular networks GPS derived position can be used to simplify the complicated hand-off protocols.

However there are some obstacles which are constraining the more wide spread usage of GPS technology. The GPS antennas must point to open sky and cannot be used inside buildings. Also, the GPS performance is affected in the midst of heavy foliage and high rise concrete structures. Some applications may also be affected by the pseudorandom noise introduced into the signals (called the SA, it is considered in section 2.2).

Chapter 2

Overview of the GPS System

The Global positioning system (GPS) is a satellite based radio navigation aid that provides time and three dimensional position and velocity information to suitably equipped users anywhere on or near the surface of the earth and sometimes off the earth. The NAVSTAR Global positioning system or the GPS as it is simply called was designed and deployed by the US Department of defence (DOD). While there are many thousands of civil users of GPS worldwide, it is still funded and operated by the US military. It was designed for 24 hour availability in all weather conditions. The GPS satellite constellation was formally declared as having met the requirements for full operational capability (FOC) on April 27, 1995. The requirements include 24 operational satellites (Block II/IIA) functioning in the assigned orbits and the system performance having met the operational military requirements. Prior to FOC, an initial operational capability (IOC) was declared on December 8, 1993.

2.1 Principle of GPS Operation

The GPS utilizes the concept of time-of-arrival (TOA) ranging to determine the user coordinates. This concept involves measuring the time of propagation of a signal, transmitted from an emitter at a known location. The time of propagation multiplied by the speed of the signal gives the distance of the receiver from the transmitter, say x_1 . In three dimensional space the receiver can be anywhere on a sphere of radius x_1 with the transmitter at the center. By simultaneously computing the distance from three different transmitters, the

receiver's position will be given as one of the intersection points of the three spheres. The above discussion assumes accurate and precisely synchronized clocks at the transmitter and the receiver. In the GPS system, the satellites transmit a broadcast code. From this, the receiver can determine the time the code was transmitted and can estimate the position of the particular satellite at the above instant. The orbiting satellites maintain extremely accurate atomic clocks, but if GPS receivers had similar atomic clocks, they would be too expensive for normal use. It turns out that GPS receivers can use inexpensive quartz clocks and yet come up with accurate position fixes. The key lies in making the distance measurements to one more satellite. These distance measurements, called the pseudoranges, include errors from sources like receiver clock offset, atmospheric signal delay, and others. With four or more pseudoranges, the receiver has enough information to calculate its latitude, longitude, altitude, and clock offset. Thus, the GPS receivers, in addition to geographical position, can give an accurate time and frequency reference signal as a byproduct.

GPS provides two levels of services: the Standard positioning service (SPS) and the Precise positioning service (PPS) [1]. The SPS is a positioning and timing service which is available to all the users worldwide with no direct charge. The SPS provides a predictable positioning accuracy of 100 meters (95%) horizontally and 156 meters (95%) vertically. The time transfer accuracy w.r.t. UTC is within 340 nanoseconds (95%). The SPS users use the GPS L1 frequency. The GPS signal structure is considered in the following sections. The other service provided is the Precise positioning service (PPS), a highly accurate positioning and timing service. It is available only to US government authorized users like the military. It is denied to the general users through cryptography. It provides a predictable positioning accuracy of at least 22m (95%) horizontally and 27.7 m (95%) vertically. The accuracy of the time disseminated will be within 200 nanoseconds (95%) with reference to UTC. The PPS users use both the GPS L1 and L2 frequencies.

2.2 GPS Signal Characteristics

The GPS ranging signal is broadcast on two L-band frequencies: a primary signal at 1575.42 MHz (L1) and a secondary signal at 1227.6 MHz (L2). Both the signals at L1 and L2 frequencies can support two modulations at the same time through a technique called phase quadrature. The present GPS system incorporates two modulations on the higher frequency (L1) but only a single protected modulation on L2. The two modulations are:

1. *C/A or Clear Acquisition code* The C/A code modulates the L1 carrier. It is a short PRN (pseudo-random noise) code with a chip rate of 1.023 MHz and a period of one millisecond. This is the principal civilian ranging signal and it is always broadcast in the "clear" (i.e. not encrypted). The usage of this signal comes under the category of Standard positioning service (SPS). For those users who are authorized to use the P code, the C/A code is used initially to acquire the much longer P code. The C/A code is available only on the L1 frequency.
2. *P or Precise code* It is a very long code (actually segments of a 200-day code are employed) that is broadcast at ten times the rate of C/A, 10.23 MHz. It has a period of 7 days and repeats every midnight Saturday/Sunday. This code acts as a source of precise ranging signal because of its higher modulation bandwidth. This reduces the noise in the received signal but will not improve the inaccuracies caused by biases. The US military has encrypted the P signal. The encrypted P code is termed the Y code. The P(Y) code is the basis for PPS. This is available only to the authorized users who have access to the cryptographic keys. When the P code is encrypted, it is termed Anti-Spoofing (AS) mode of operation. This is because the unpredictable code (the Y code as it appears to unauthorized users) cannot be spoofed or jammed. Jamming is a technique in which adversaries try to replicate the ranging codes and carrier frequencies to mislead the user.

The P code is available on both L1 and L2 while the C/A code is restricted to L1. Because of the AS mode of operation, the civilian receivers operate only at the L1 frequency. The advantage of operating at both the frequencies, as the P(Y)-code users do, is that the ionospheric delay can be calibrated. The free

electrons in the ionosphere introduce a delay in the modulation signal (PRN code). Thus the pseudoranges computed have to be corrected for this error. The delay introduced is proportional to the integrated number of free electrons along the transmission path and inversely proportional to the square of the transmission frequency (to first order). The delay is related by a scale factor to the difference in signal time of arrival for the two carrier frequencies. This relationship enables the P(Y) code receivers to make dual frequency measurements at L1 and L2 and measure the delay directly. On the other hand, the C/A code receivers, restricted to operate at one frequency, employ an ionospheric model to estimate the delay. The 8-model parameters used to calculate the correction are broadcast as part of the GPS navigation message. (In addition to the PRN codes, a 50 bps navigation message also modulates the L1 carrier. The navigation message contains sufficient information to determine the GPS satellite orbits, clock corrections, and other system parameters.) The delay as determined from the model by the C/A code receivers is comparatively less accurate than the one measured directly by P(Y) code receivers. However, codeless receivers do provide a way for the civilian users to determine more accurate ionospheric delays.

Another constraint the C/A code receivers face is what is called Selective Availability. Selective Availability, or SA, is the intentional degradation of the signals by DOD to deny full accuracy to the SPS users. It is accomplished by manipulating navigation message orbit data (epsilon) and/or satellite clock frequency (dither). A technique known as differential GPS can overcome this limitation. The above topics and other details of the GPS system are comprehensively covered in [2].

2.3 GPS System Segments

The GPS system comprises of three major segments: space, control, and user.

1 Space Segment The space segment consists of the satellites in orbit that provide the ranging signals and data messages to the user equipment. The user must make pseudorange measurements to four or more satellites simultaneously.

for real time three dimensional navigation. Thus four or more satellites must be visible from any point of the earth at any time of the day. Keeping this requirement in view, the GPS satellite constellation is designed to consist of 24 satellites. The satellites are positioned in six earth-centered orbital planes with nominally four satellites in each plane. The orbits are nearly circular and equally spaced about the equator (60 deg apart) with an inclination of 55 deg with respect to the equator. The satellites have an orbital period of one half the sidereal day (i.e. 11hr 58 min) and the semimajor axis is of 26561.75 km. Since these satellites are in a 12 hour sidereal time orbit, they produce a ground trace (projection on earth's surface) which repeats over and over once every day. The above constellation provides the user with, between five to eight visible satellites, from any point of the earth.

Several notations are in use to refer to a particular satellite in orbit. A satellite may be characterized by a space vehicle number (SVN) which is assigned to it by the US airforce, or by the PRN codes that it generates. These PRN code generators are configured uniquely on each satellite producing distinct versions of C/A and P(Y) codes for each satellite. These satellites are also grouped into several blocks depending upon with what phase of development they have been associated with i.e., a block of satellites are similar in design respects and are launched more or less together. The initial concept validation satellites are called Block I satellites. The present operational GPS satellites are designated into Block II (SVN 13-21), Block IIA (SVN 22-40) and Block IIR (SVN 41-62). At the present, Block IIR satellites called the replenishment satellites, are being launched. The design life of a Block IIR satellite is about 7.8 years and these will be the principal ranging satellites in the coming years. As on April'98 the current GPS constellation consists of 27 Block II/IIA/IIR satellites [1].

2 Control Segment The control segment oversees the functioning of the GPS system and is responsible for maintaining its correct operation. It maintains the satellites in their assigned orbits, carrying about small orbital corrections if necessary (called station keeping). It also monitors the satellite subsystem health and activates spare satellites if available and necessary. It also controls

the SA and AS modes of operation. However the most important operation the control segment carries out is updating each satellite's clock ephemeris almanac and other parameters in the navigation message. The ephemeris parameters are a precise fit to the GPS satellite orbits and help the receiver to determine the exact orbital coordinates of the satellite at a particular instant. The almanac is a reduced precision subset of the ephemeris parameters. Almanac data is used to predict the approximate position of all the satellites and aids the receiver in acquiring the satellite signals.

The control segment consists of the three components: the monitor stations, the master control station (MCS) and the ground antennas. There are five monitor stations at Hawaii, Colorado Springs, Ascension Island, Diego Garcia in the Indian Ocean and at Kwajalein island in the West Pacific which form the data collection component of the control segment. The monitor stations perform ranging measurements to all satellites in view with a dual frequency (L1/L2) GPS receiver. This raw data, along with the received navigation messages are passed onto the master control station located at Falcon airforce base in Colorado. A lot of processing takes place on the collected data to form estimates of the GPS satellite clock ephemeris and almanac. The collected pseudoranges and deltaranges are first corrected for ionospheric/tropospheric delays and then processed by an epoch-state Kalman filter to form a precise satellite ephemeris and clock offset solution. The satellite position and clock corrections are predicted forward in time using precise models that are valid over a particular prediction interval. The prediction interval is subdivided into either 4 or 6 hour time intervals and the satellite position data is transformed into 15 curve fitting orbital elements. The almanac (7 of the 15 ephemeris orbital parameters) and clock data are also determined from this accurately predicted data. This updated data is transmitted to the ground antennas and from there to the satellites navigation processor via a S band data communication uplink.

3 User Segment The GPS user segment consists of the GPS receivers and the user community. GPS receivers determine the user's position, velocity and time estimates (PVT) from SV signals. The receiver structure is considered later.

2.4 Carrier Modulation

The GPS system employs Direct sequence spread spectrum multiple access techniques (a type of CDMA). So all the GPS space vehicles (SVs) can transmit on the same carrier frequencies L1 and L2, but their signals do not interfere significantly with each other. The carrier frequencies are modulated by spread spectrum codes with a unique PRN sequence associated with each SV and by the navigation data message. Fig. 2.1 is a representative view of how the GPS signals are generated [3].

All the clock rates required for the generation of PRN codes, radio frequency carriers and for the navigation data stream are derived from a single reference frequency and are thus coherently related. The nominal reference frequency as it appears to an observer on the ground is 10.23 MHz. To compensate for relativistic effects (the satellite clock is affected by both special and general relativity) the satellite clock is offset by a factor of $\Delta f / f = -4.467 \times 10^{-10}$, giving a Δf of -4.57 mHz prior to launch. To an observer on the satellite vehicle, the satellite's clock frequency is seen as

$$f_0 = 10.22999999543 \text{ MHz}$$

Both the carrier frequencies are a multiple of the 10.23 MHz master clock. In particular

$$L1 = 154 \times 10.23 = 1575.42 \text{ MHz}$$

$$L2 = 120 \times 10.23 = 1227.60 \text{ MHz}$$

The L1 carrier is modulated by both the PRN codes but in phase quadrature with each other. There is a 90° phase shift between the C/A code + data modulation and P(Y) code + data modulation. The L2 frequency is modulated by only one PRN code at a time. The control segment controls the switch output which can be P(Y) code + data or P(Y) code or C/A code + data. Normal operation provides P or Y code modulation on L2. When the AS mode of operation is activated, the P code is encrypted to form the Y code which also has the same chip rate. The SPS users are thus denied access to the precision code and so civilian users operate at L1 only. Fig. 2.1 also shows the incorporated dither frequency when SA is activated. Both the C/A code and the P(Y) code as well as the L1 and L2 carrier frequencies

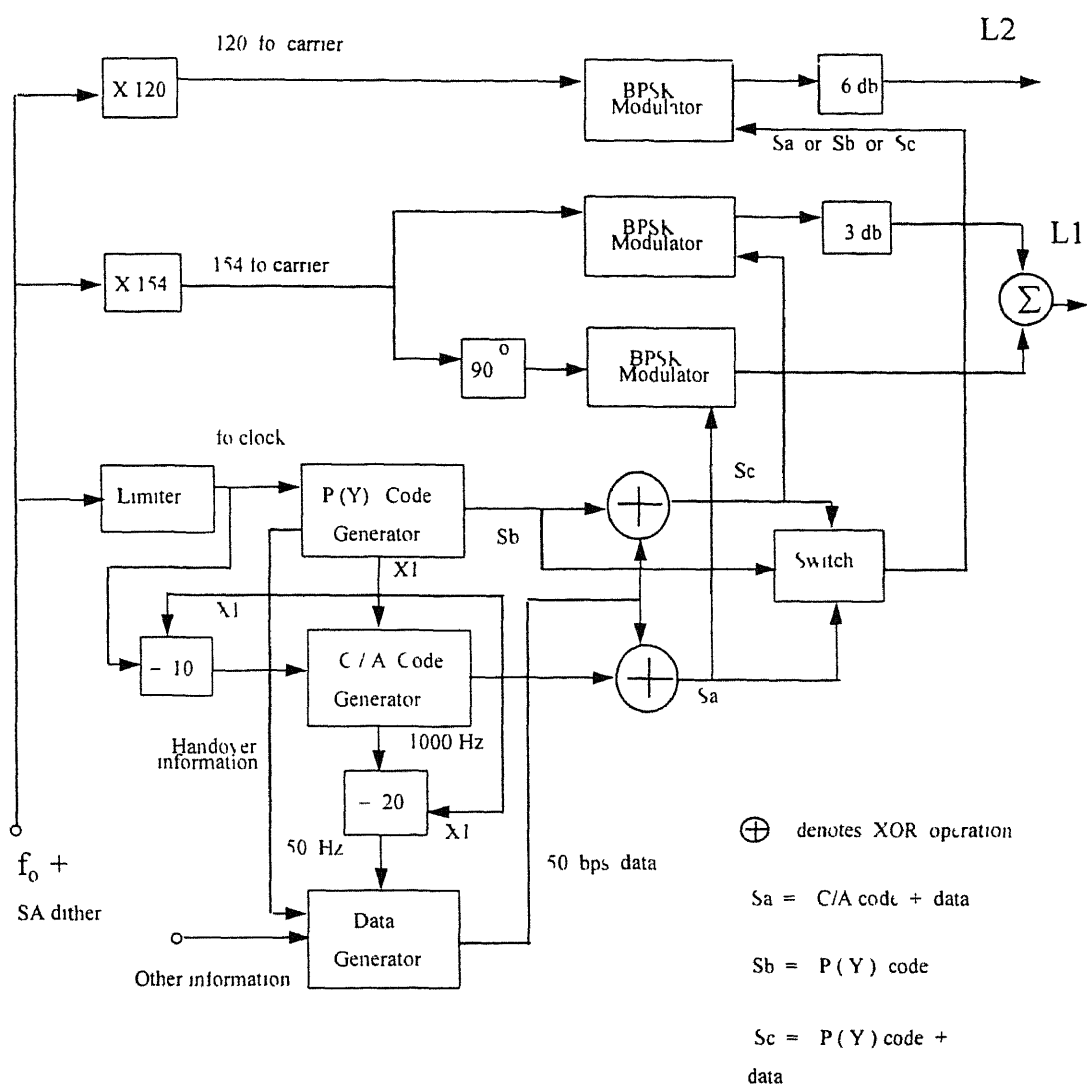


Fig 2.1 GPS signal structure

are subjected to the encrypted dither frequency of SA. The SA error can be removed by the PPS user but cannot be corrected by SPS user. In the SV transmitted signal, the C/A code strength is nominally 3dB stronger than the P code on L1. Both the C/A and P codes are of a class called product codes, each is the product of two different code generators clocked at the same rate where the delay between the two code generators defines the particular satellite code. The code generator details can be found in [2] and [3]. The C/A code is a relatively short code with a period of 1023 chips or 1 ms duration. These short duration C/A codes have been selected so as to permit rapid acquisition of the signal. The C/A codes belong to a family of codes known as Gold codes, which

are formed by the product of two equal period 1023 chip PN codes $G_1(t)$ and $G_2(t)$. The product code $G(t)$ will also be of duration 1023 chips and is represented as

$$G(t) = G_1(t) \cdot G_2(t + n_i \cdot 10 T_c)$$

where $10T_c$ is the C/A code chip period and n_i will determine the phase offset in chips between G_1 and G_2 and this delay is unique to each SV.

The P codes are also the product of two PN codes and are represented as

$$P(t) = X_1(t) \cdot X_2(t + n_i T_c)$$

Here T_c is the P code chip period. The value of n_i used is slightly different from the above one but both of them are related to the PRN value associated with the SV. The X_1 sequence has a period of 1.5 seconds or 15345000 chips while the X_2 sequence is 37 chips longer. Since the periods of X_1 and X_2 are relatively prime to each other, the P code has a period equal to

$$\begin{aligned} &= (15345000) \times (15345037) \\ &\cong 2.3 \times 10^{14} \text{ chips} \end{aligned}$$

Thus if the P code were allowed to continue without being reset, each P code would continue without repetition for slightly more than 38 weeks. This overall period has been subdivided so that each SV gets a 1 week period code which is non-overlapping with that of any other satellite. It can support a maximum number of 37 possible GPS satellites or ground transmitters (pseudolites). The X_1 and X_2 sequences are reset at the start of a week (Saturday/sunday midnight) and both of them begin the week at the same epoch time. Fig 2.1 also shows the 50 bps navigation message data which is uploaded to the satellites by the GPS control segment. The navigation data provides several key inputs that enable the user to obtain a satisfactory navigation or time transfer solution. The 50 bps data stream is modulo-2 added to both the C/A and P(Y) codes on L1 frequency, but may or may not be carried on L2 depending on the satellite mode. The data is grouped into frames and subframes.

A simplified GPS frame structure is shown in Fig 2.2. A frame consists of 5 subframes while 25 frames form a superframe. A subframe consists of ten 30-bit words. A subframe thus lasts for 6 seconds in duration, a frame for 30 seconds, and a superframe for 12.5 minutes.

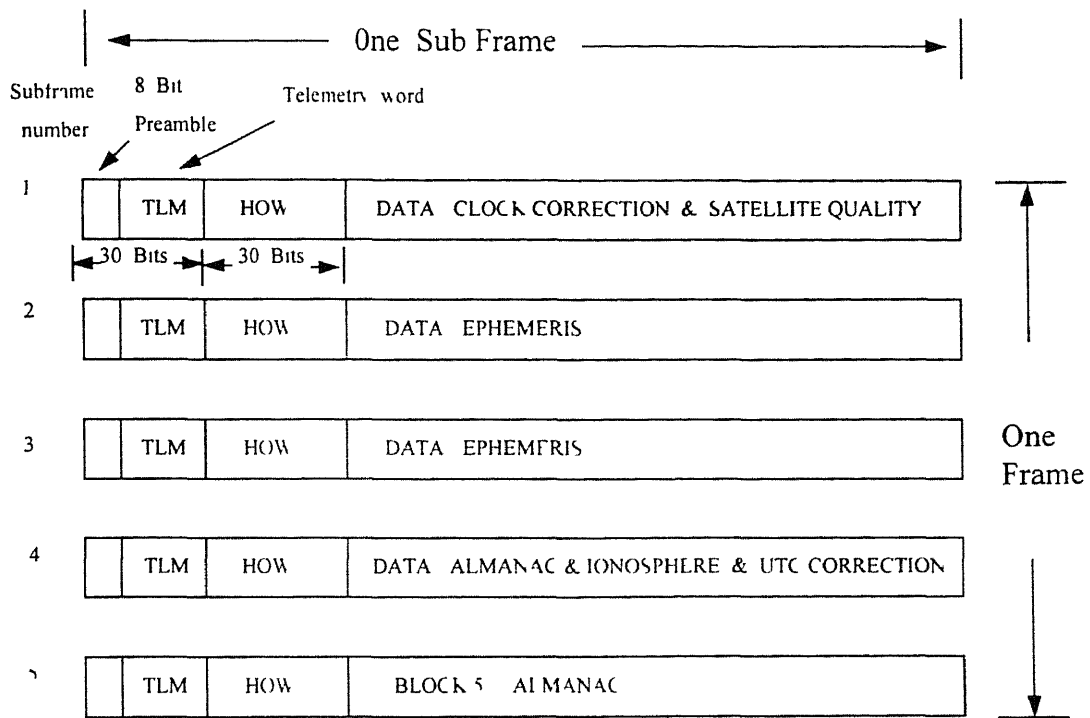


Fig 2.2 GPS Frame and Subframe format

In the figure the 8-bit preamble shown of each subframe is used for synchronization purposes. A long period code such as the P code is difficult to acquire without any acquisition aids. The Hand-over-word (HOW) transmitted keeps track of the number of P(Y) code 1.5 second duration X1 epochs thus far in the week. Once the receiver has knowledge of the HOW information (it gets this by tracking the C/A code) it can acquire the P(Y) code in the next subframe. The ephemeris provides the receiver the precise position of the satellite at the time of transmission while the almanac gives the approximate satellite ephemeris, clock correction and satellite health for the entire GPS constellation. This information helps the receiver in selecting the best set of satellites (for e.g. appropriately low GDOP) for the navigation and time solution.

2.5 GPS Receiver Structure

The GPS receiver processes the L band signals transmitted from the satellites to determine the user PVT. A functional block diagram of a generic

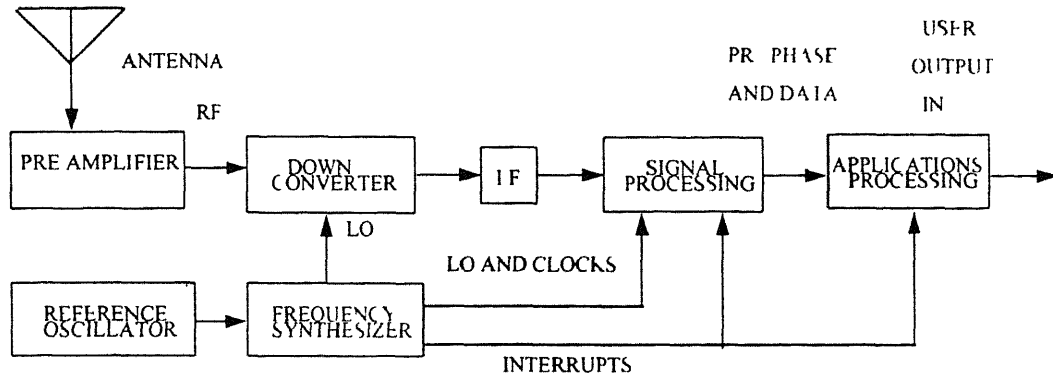


Fig 2 3 GPS Receiver Functional Block Diagram

GPS receiver is shown in fig 2 3 Satellite signals are received via the antenna which is right hand circularly polarized and provides near hemispherical coverage GPS receivers that track P(Y) code on both L1 and L2 need to support 20 46 MHz bandwidths on both the frequencies If the receiver uses C/A code only, the antenna and receiver must support a bandwidth of atleast 2 046 MHz Antenna design selection is made keeping in view the intended applications and the designs vary from helical coils to thin microstrip patch antennas The preamplifier section rejects out of band signals and contains a LNA for signal amplification The reference oscillator provides the time and frequency reference for the receiver The reference oscillator plays a key role as all the GPS receiver measurements are based on the time-of-arrival of pseudorandom (PRN) code phase and received carrier phase and frequency information The frequency synthesizer makes use of the reference oscillator output and from it derives a number of local oscillators and clocks used by various stages of the receiver These local clocks comprise the receiver's time base The downconverter converts the RF inputs to IF, which is easier to process The IF section provides further filtering and maintains appropriate amplitude levels for the signal plus noise to facilitate further signal processing

The signal processing block constitutes the core of a GPS receiver and carries out a whole lot of functions such as

- 1 Splitting the signal into multiple channels for signal processing of multiple satellites simultaneously

- 2 Generating the reference PRN codes of the signals
- 3 Acquiring the satellite signals
- 4 Tracking the code and carrier of the satellite signals
- 5 Demodulating the received signals and extracting the system data
- 6 Extracting the PRN code phase (pseudorange) measurements
- 7 Extracting carrier frequency (pseudorange rate) and carrier phase (delta pseudorange) measurements from the carrier of the satellite signals
- 8 Extracting the signal-to-noise ratio SNR from the satellite signals
- 9 Estimating the GPS system time

The outputs of the signal processing function are pseudoranges, delta pseudoranges and pseudorange rates, signal-to-noise ratios local receiver time tags and GPS system data for all the satellites being tracked all of which are provided to the application processing function. The applications processing block varies depending on the type of end requirements such as performing time transfer navigation or differential surveying. It controls the signal processing functions for providing it the appropriate measurements and data.

2.6 Receiver Classification

The classification of GPS receivers can be done on the basis of many aspects. The receivers can be classified on the basis of the receiver architecture, the method of operation or on the type of applications they have been designed for.

Based on Architecture

- 1 *Sequential Receivers* The sequential receiver uses one or two hardware channels to sequentially provide individual satellite observations. These receivers are the cheapest but provide the poorest time-to-first-fix and cannot track the satellites while moving at high speeds.
- 2 *Continuous Receivers* The continuous receiver has sufficient number of dedicated hardware radio channels to continuously track the satellites and provide individual satellite observations, thus giving it a very high performance. A minimum of four channels are required for continuous

operation. A five channel receiver can view four satellites and use the fifth for reading the navigation message thus continuously updating the receiver's database of the satellite orbital and other parameters. Receivers with more than five channels can track all the visible satellites and put them in reserve in case one of the four satellites, presently being used for the navigation and time solution, is lost for any reason.

- 3 *Multiplex receivers* The multiplex receiver acts like a sequential receiver in that it switches between satellites being tracked. However it does this at a fast sample rate (approximately 50 Hz) and can track more satellites than a sequential one. Its performance is still lower than that of a continuous receiver because it cannot integrate all of the satellites' transmitted spread spectrum power.

Based on the Method of Operation

- 1 *Code Correlation* These receivers arrive at the PVT solution by processing the PRN codes transmitted by the satellites. The receiver creates an internal replica of the known PRN modulation sequence and adjusts the internal epoch until it exactly matches the received signal in delay. This matching is performed by cross-correlating the received and internal signals and finding the start time that maximizes the correlator output. The pseudoranges are then determined from a knowledge of the receiver's clock and transmission times. While they have the advantage of low cost, they only provide a moderate accuracy.
- 2 *Carrier Phase* These receivers reconstruct the GPS radio frequency carrier and use the sinusoid as a ranging signal. These carrier phase measurements are used to arrive at the PVT solution. They do not need to decode the SV transmitted signals except for locating the satellites. Some such receivers may have no code reception capability at all, in which case the receiver must be pre-loaded with that data from another source. These measurements are very precise but the accuracy may be limited by the difficulty in resolving which cycle is being received. These receivers have been used in applications like surveying where centimeter level accuracies have been

attained by differential corrections. However their use is prohibited by their high cost.

Based on Applications

The GPS receivers are designed keeping the end applications for which they are intended to be used for, in view. As the number of areas in which GPS can be put to use are increasing day by day, more and more types of receivers are being conceived and designed. A few of the specialized ones are noted below.

- 1 *Aviation* Airborne GPS receivers are generally used for navigation and attitude determination. High-end GPS units are being built into aircraft for automated landing. These receivers may use multiple antennas, the relative placement of which must be known.
- 2 *Space-borne* These receivers are used on satellites both for navigation and attitude determination. These are radiation hardened and have special features to allow them to operate satisfactorily in the face of high relative velocities experienced by the orbital spacecraft.
- 3 *Mapping* These receivers are optimized for collecting data to be exported to an external database. These often support DGPS operations and are made to be highly portable. Often they will have an attached computer dedicated for data collection.
- 4 *Surveying* These receivers are used by surveyors to derive "accurate measurements" rather than "position", i.e. it is the relative relationship between two receivers which is more important and from this an absolute position may be derived if necessary. Surveying receivers are generally capable of the highest accuracies and cost a fortune.
- 5 *Timing and Frequency* These types of receivers are intended to act as a time and frequency reference. Position is of secondary importance here and is often ignored. Some of the receivers may be furnished with a 1 pulse-per-second output while others may have attached rubidium or cesium atomic clocks for improved short term and long term stability. High end timing units are used to synchronize digital telecommunication networks.

2.7 GPS and Time

The GPS system is based on the reception of radio signals synchronized with atomic clocks. Thus the GPS receivers, in addition to the geographical position of the receiver site, can give an accurate synchronization signal (time and frequency) as a byproduct. The GPS has been shown to be an outstanding tool for the dissemination of UTC (Universal coordinated time). The UTC is a composite time scale used as a time reference all over the world. It incorporates the uniformity of an atomic time scale but at the same time it is in step with the non-uniform time scale determined on the basis of Earth's rotation and revolution. To keep it synchronized with the solar time occasional leap seconds are introduced into the UTC time scale.

Each block II/II A satellite contains two cesium (Cs) and two rubidium (Rb) atomic clocks. Each SV is thus provided with precise frequency references and also maintains sufficient redundancy. GPS time is a time scale established by the control segment and referenced to UTC (USNO). UTC(USNO) is the UTC time scale as maintained by the US naval observatory (USNO). The USNO monitors the timing of the GPS and provides a reliable and stable time reference. The GPS system time is a paper time scale based on statistical processing of data obtained from all the monitor stations and the satellite operational frequency standards.

GPS time is a continuous time, i.e. it is not adjusted and therefore is offset from UTC by an integer number of seconds (If it were adjusted by leap second insertions, it would throw the GPS P(Y) code receivers using the system out of lock). The number remains constant until the next leap second occurs. This offset is given in the navigation message and the receiver applies the correction automatically. As of July 1, 1997, when the last leap second insertion took place, GPS time is ahead of UTC by 12 seconds. The GPS time is referenced to the Master clock (MC) at the USNO and steered to UTC (USNO) from which system time will not deviate by more than 1 microsecond (modulo 1s). The exact difference is contained in the navigation message in the form of two constants, A_0 and A_1 , giving the time difference and rate of system time against UTC (USNO, MC). GPS time is steered to the USNO reference on a

daily basis and during the last several years it has been maintained within a few hundred nanoseconds. The rate of steer being applied is $\pm 1.0 \times 10^{-19}$ seconds per second squared [1].

How the receiver computes UTC (USNO) can be seen in the following equation

$$\text{UTC (USNO)} = T_{\text{rec}} + T_u + T_n$$

T_n is the number of leap seconds that GPS time lags behind the UTC. T_{rec} is the receiver clock time. T_u is the bias from the GPS system time which the receiver determines from the PVT navigation solution. Before computing the solution, the receiver also adds a correction, determined from the navigation message, to the readings of the on-board satellite clocks. Thus the user can refer time via the GPS to the UTC (USNO) directly, the accuracy of the time being within 200 nanoseconds (PPS requirement) or within 340 nanoseconds for the SPS users.

2.8 GPS Error Sources

GPS errors are a combination of noise, bias, and operational errors. Noise errors are the combined effect of PRN code noise (around 1 meter) and noise within the receiver (around 1 meter). Bias errors can result from Selective Availability (SA). The SA bias on each satellite signal is different, and so the resulting navigation solution error is a function of the combined SA bias of all the SVs used in the solution. However, the SA error is spatially correlated, i.e., the error at one location bears a relation to the error at a nearby location. This property can be used to remove SA errors through differential techniques. Other bias errors may include ephemeris data errors and uncorrected SV clock errors. The C/A code receivers estimate the tropospheric and ionospheric delays through pre-determined models. The models are not so accurate, and residual errors remain. Another kind of bias error occurs due to multipath signals. Multipath is caused by reflected signals from surfaces near the receiver that can interfere with or be mistaken for the signal that follows the straight line path from the satellite. Multipath is difficult to detect and often hard to avoid. The operational error sources include receiver hardware/software failures, incorrect

geodetic selection by the user or control segment failures

The position errors that result from all these measurement errors depends on the user/satellite relative geometry. This leads to the concept of Geometric Dilution of Precision (GDOP). The DOP parameters are geometry factors that relate parameters of the user position and time bias errors to those of the pseudorange errors. Under the assumption of uniform, uncorrelated zero-mean ranging error statistics this can be expressed as

$$\text{RMS position error} = (\text{Geometric dilution}) * (\text{rms ranging error})$$

Thus lower the GDOP, the better. Geometrically speaking, the volume of the shape formed by the unit vectors from the receiver to the SVs used in the position solution is inversely proportional to GDOP. Greater is the volume, smaller is the value of GDOP and better is the resulting position accuracy. This results when the angles from the receiver to SVs are different. On the other hand, smaller is the volume, larger is the value of GDOP and poorer is the resulting position accuracy. This results when the angles from the receiver to selected set of SVs are all similar. To characterize the accuracy of various components of the position and time solution, various other DOP parameters are commonly used. These are

PDOP - Positional dilution of precision (sometimes called the spherical DOP (3-dimensional position accuracy))

HDOP - Horizontal dilution of precision (Latitude, Longitude)

VDOP - Vertical dilution of precision (height)

TDOP - Time dilution of precision (Time)

While each of these GDOP terms can be individually computed, they are formed from covariances and so are not independent of each other. A high TDOP, for example, will cause receiver clock errors which will eventually result in increased position errors.

2.9 Differential GPS

Differential GPS (DGPS) is a technique to improve standalone GPS accuracy by removing the common, i.e. correlated, errors from two or more receivers viewing the same set of satellites. DGPS requires a reference station

which has a precise knowledge of the coordinates of its antenna's phase center. The reference station makes the usual code-based GPS pseudorange measurements but knowing its position it can determine the biases in the measurements. These biases are computed for each of the set of all visible satellites by differencing the pseudorange measurements and the satellite-to-reference station geometric range. For real time applications the reference station transmits these differential corrections to all the users in a given local area. The users incorporate these corrections before processing their individual pseudoranges and thus improve the overall accuracy of their navigation and time solution. DGPS offers accuracies of a few meters (better than the PPS service) and facilitates detection of erroneous signals from the satellites. If the receiver is close to the reference station the error components attributable to the space and control segments may be entirely removed while those contributed by the user segment may be substantially removed. Some applications like surveying utilize carrier-based DGPS to obtain highly precise results.

2.10 Accords GPS Receiver

We have used the Accords GPSR 2000 receiver in carrying out this thesis work. The receiver is based on the Analog Device's floating point family of DSP's and offers a software-intensive solution for the signal correlation, positioning and navigation and time determination. The receiver can be connected to the serial port of a computer via a RS232c link through which the receiver can be controlled and its data can be monitored. The receiver's functional block diagram is as shown. The ADSP 21062 (SHARC - Super Harvard architecture computer) processor is a 32-bit floating point DSP with on-chip SRAM and provides a programmable platform for the receiver. GPSR 2000 is designed to simultaneously acquire and track 8 satellite signals at a time and to compute the user's position, velocity and time.

The receiver is designed to operate on the C/A ranging signal. So the antenna, which has a built-in low noise amplifier, operates at the 1575.42 MHz

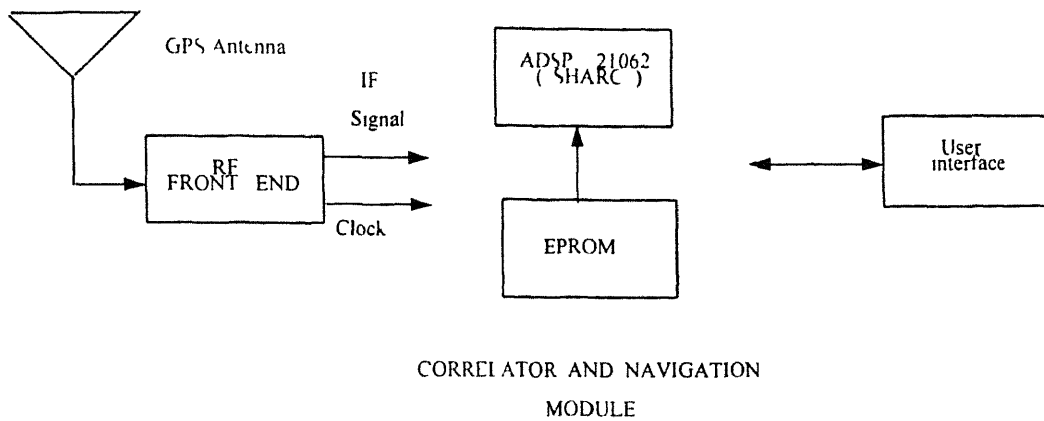


Fig 2 4 GPSR 2000 Receiver Model

range. It has a near hemispherical radiation pattern. The satellite signal at the antenna is very weak (-130 dbm) with an SNR of less than -16dB. In the RF section, the incoming signal is first amplified and processed by a Bandpass filter operating at 1575.42 MHz and later downconverted to a low IF in three mixer stages. The IF is then digitized and fed to the correlator and navigation module.

The correlator and navigation module store the incoming digitized IF samples in a buffer and carry out a lot of intense signal processing operations to arrive at the PVT solution. The operations include Down-conversion of IF signals to baseband, generation of local carriers and C/A codes for all the satellites, carrying out the appropriate correlations and determining the pseudoranges, satellite selection and satellite database management and processing of the measured data to give an accurate user position and time. All the 8 channels are processed parallelly. The user can monitor and control the receiver through the user interface. The receiver can also be given an initial estimate of the user's position and time to reduce the time-to-first fix. The GPSR 2000 receiver's output formats which have been used in this thesis work are described in brief in Appendix A. More complete details can be found in [4].

Chapter 3

Network Time Synchronization

An often overlooked aspect but albeit an important one in computer network design is a high quality time-of-day clock capable of generating accurate timestamps with small residual errors compared to intrinsic one way network delays. Such a service would be useful for tracing the progress of complex transactions, synchronizing distributed data bases, monitoring network performance and isolating problems. The underlying problem, however is not setting just one clock on one machine, but instead keeping an entire network of machines synchronized with respect to a global reliable time source and with respect to each other.

Just as events in our personal lives are controlled by time, so are events in networked computer systems. Accurate time-keeping manifests in many day-to-day functions as well as in issues critical to system administration. A few of them are dwelt upon below.

- Many operating systems, UNIX for instance retains the latest version of a file involved in a transfer. If a file is taken from a server modified and appended with an incorrect time, the modified file can be lost. Utilities like “make” consult file modification times, so offset clocks may result in erratic actions.
- Complex simulation tasks can be divided amongst multiple processors. The correct sequence of events is maintained by assuring that all platforms have the same time.
- Several network services expect that system clocks on client machines are



synchronized with the clocks on server machines. For example, using a secure RPC service such as secure NFS or NIS+, accurate time management is a must to ensure that timestamp verifiers generated by a client are accepted by the server. Client clocks that stray about result in rejected requests.

- Many network security systems are based on accurate time tagging at each end of a communication path. Some measure time of transit to reject transactions with excessive delays and others issue authentication tickets that are only valid within a tightly controlled time window. For example, the “Kerberos” [6], a widely used network authentication protocol, assumes that all the clocks are synchronized to within several minutes.
- Attempts to correlate network activity, performance data, and system log messages from multiple hosts require a single, global timepiece for the network. Tracking a performance problem across several machines is feasible only if we can lay out the events on a single absolute timescale.
- Time synchronization issues play an important role in networks supporting real-time services such as distributed multimedia conferencing and in other applications involving distributed resources like financial networks and others.

Thus, as an increasing number of network services depend on timestamps, verification and sequencing, it becomes important to look at ways to keep the widely distributed clocks maintain time within acceptable tolerances.

Many protocols have been discussed in the literature to synchronize and to disseminate time. Some of the mechanisms specified include ICMP timestamp message, Time protocol, and the Daytime Protocol. Both the Time and Daytime protocols do not support any kind of compensation for the transmission delay between the client and the server. For this reason, they are only accurate enough for computers on the same network. The Unix 4.3 BSD time daemon `timed` uses a single master-time daemon to measure offsets of a number of slave hosts and send periodic corrections to them. In this model, the master is determined using an election algorithm which requires the ability to broadcast. Thus, the types of networks on which this can be used are limited.

Included in the Berkeley r-commands is `rdate`, the equivalent of the date time setting command that takes its input from another host on the network. The result will be system clocks that are synchronized to a common source. In addition to the accuracy being limited to a few seconds, the client clocks will only be as accurate as the server singled out as a time source. Users may also see time discontinuities. What we want is a time-management protocol that adjusts the clock slew rate, creating gentle shifts in time and uses a group of trusted servers to produce an accurate stable time base. It's called the Network Time protocol or NTP. Before going through the NTP in detail let us first look at how the computer clocks are organized.

3.1 Computer clocks

A local clock is used in each computer in order to maintain the time. It includes an oscillator, clock counter and software support to provide the time in some format to the operating system and client processes. Most of the Unix clock models require a periodic interrupt produced by the hardware frequency source in the 100-1000 Hz range. For instance Sun Sparcs and in Intel 80386s the interrupts occur every 10 ms. Each interrupt causes an increment called tick to be added to the software clock counter. The value of the increment is chosen so that the counter plus an initial offset established by the `settimeofday()` call is equal to the time of day in seconds and microseconds [7]. On Unix systems the time is stored as a 64-bit value representing the number of seconds elapsed since January 1, 1970. The time counter is relative to Universal Coordinated time (UTC) and is converted into local time using the information in the `timezone` file `/etc/timezone`.

Most computers use an uncompensated crystal oscillator as the hardware source for generating interrupts. These exhibit some natural drift which is difficult to characterize. It may be as high as to make the clock lose a few seconds a day. Another common source of time errors is when the kernel loses an interrupt. The hardware interrupts only cause the kernel to increment a variable that keeps track of the time. During moments of high system load, the kernel may be too busy for the period between two consecutive interrupts, i.e., as

the operating system clock is interrupt driven it may lose time during operations that span more than one clock interrupt period

3 2 The Network Time Protocol

NTP is a protocol built on top of TCP/IP that assures accurate local timekeeping with reference to radio, atomic or other clocks located on the Internet. It has been engineered to maintain accurate and reliable time in the face of typical internet conditions characterized by multiple gateways, variable delays and unreliable networks. The NTP is used by Internet Time servers and their clients to synchronize clocks and as well as to automatically organize and maintain the time synchronization subnet itself. The NTP provides client accuracies typically within a millisecond on LANs and upto a few tens of milliseconds on WANs relative to a primary server synchronized to Coordinated Universal time (UTC) via a GPS receiver for example.

The network time protocol is now established as an Internet standard protocol and is presently in its third version [8] constitutes a formal specification of the protocol. NTP is built on the Internet protocol (IP) and User datagram protocol (UDP) which provides a connectionless transport mechanism. The approach used by NTP to achieve reliable time synchronization from a set of possibly unreliable remote time servers is somewhat different than other protocols. In particular, NTP does not attempt to synchronize clocks to each other. Rather, each server attempts to synchronize to UTC using the best available source and available transmission paths to that source. NTP operates on the premise that there is one true standard time, and that if several servers which claim synchronization to standard time disagree about what that time is then one or more of them must be broken from the synchronization subnet. The NTP expects that the time being distributed from the root of the synchronization subnet will be derived from some external source of UTC. However in networks isolated from UTC sources NTP can still be used to synchronize and distribute time. This is done by nominating one or other machines with stable clocks as phantom UTC sources. The following subsections describe in brief the working of NTP. The description is based on the one given in [9]

3 3 System Architecture

The NTP system consists of a network of primary and secondary time servers, clients and interconnecting transmission paths. Time is distributed through a hierarchically organized subnet. The time server is operating at various levels of the subnet exchange periodic messages containing precision time stamps to adjust local oscillator phase and frequency. A Primary time server is the one at the root of the synchronization subnet and has access to some external time source. They are directly synchronized to a primary reference source which can be a calibrated atomic clock, a GPS receiver, a time code receiver or others. More information on the various primary time servers and their synchronization accuracy can be found in [10]. A secondary time server derives synchronization either directly from a primary server or via other secondary servers which in turn may be synchronized to a primary server directly or via others. Thus the actual synchronization paths assume a hierarchical configuration with the primary reference sources at the root and servers of decreasing accuracy at increasing levels towards the leaves.

This is analogous to some digital telephone networks where to synchronize the various switches with a master reference frequency a hierarchical master-slave network configuration is used. Borrowing the telephony conventions the NTP subnet servers are designated by a number called the stratum. The stratum is a measure of each server's accuracy and indicates how far away from an external source of UTC it is operating at. A primary time server is designated as a stratum-1 server. A stratum-2 server is one which is currently obtaining time from a stratum-1 server. A stratum-3 server gets its time from a stratum-2 server and so on. As the stratum increases from one, the accuracies achievable will degrade depending on the network paths and local clock stabilities. The number of strata are limited to 15 to avoid long lived synchronization loops.

A schematic subnet is shown in fig 3.1 which illustrates the issues considered. The nodes represent subnet servers. As shown they are characterized by stratum numbers which are taken as the number of hops to the root. The heavy lines represent the active synchronization paths and also show the

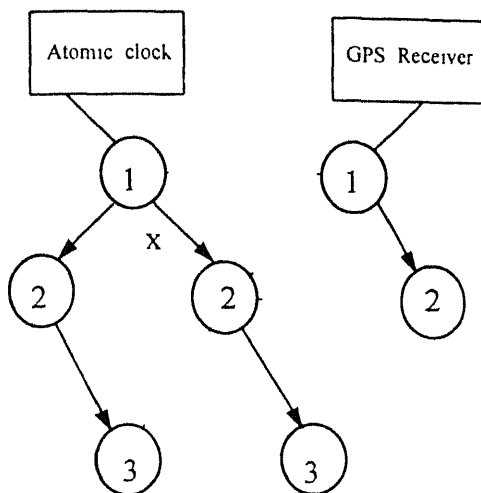


Fig 3 1 Synchronization subnet

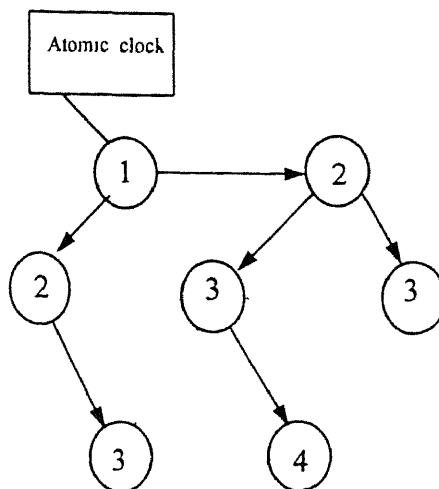


Fig 3 2 Modified Synchronization subnet

direction of timing flow The light lines represent the back up synchronization paths where NTP packets are exchanged but those servers are not necessarily used in adjusting the local clock Two stratum-1 servers are shown, one with a GPS clock and the other obtains its time from a calibrated atomic clock (for e g it could be the one maintained by a National standards laboratory) The atomic clock and the GPS receiver are considered to be at stratum zero Fig 3 2 shows the case when the network path marked 'x' is down and the GPS receiver malfunctions The subnet has automatically reconfigured itself to the one shown

The actual synchronization subnet is organized using a variant of the Bellman Ford distributed routing algorithm to compute the minimum weight spanning trees rooted at the primary reference sources The distance metric used by the algorithm consists of the stratum (scaled value) plus the synchronization distance The synchronization distance itself is a sum of a factor called dispersion plus one half the absolute delay (ie to the root) "Dispersion" is a measure of the time keeping quality at a particular peer Thus the synchronization path will always result in a minimum number of servers to the root, with ties resolved on the basis of maximum error

3 4 Modes of operation

When two peers exchange NTP messages, an association is said to be

formed. One or both the peers create and maintain the protocol related state variables depending on the type of association formed. The association can operate in one of the five modes: Symmetric active, Symmetric passive, Client Server and Broadcast. The mode of each server in the association pair indicates the behaviour the other server can expect from it. The various associations result from appropriate declarations in the configuration file. The configuration file and other working details are considered in the succeeding chapter.

Symmetric Active In this mode of operation the host announces its willingness to synchronize and be synchronized by the peer. A host operating in this mode sends periodic messages regardless of the reachability state or stratum of its peer. Symmetric active mode is intended for use by the time servers operating near the higher end of the synchronization subnet.

Symmetric Passive This type of association is normally created on the receipt of a message from a peer operating in the symmetric active mode. It persists only as long as the peer is reachable and operating at a stratum level less than or equal to the host; otherwise the association is dissolved. However it will persist till at least one message has been sent. In this mode also, the host announces its willingness to synchronize and be synchronized by the peer.

Client A host operating in this mode sends periodic messages regardless of the reachability state or the stratum of its peer. In this mode the host indicates to the peer that it wishes to obtain time from it, but it is not willing to provide time to the peer. This mode is appropriate for file-server and workstation clients that do not provide synchronization to other local clients.

Server This type of association is normally created on the receipt of a client request message. In response to the client request the server merely interchanges the address fields, fills in the required timestamps and other information, recalculates the checksum and returns the message to the client. The servers need not maintain any other state information between the client requests. For e.g. a LAN Time server can operate in this mode. By operating in this mode it announces its willingness to synchronize other peers but not be synchronized by any of them.

Broadcast The broadcast/multicast mode of operation is appropriate for use

on high speed LANs where a number of workstations have to be synchronized but the accuracy requirements are modest. One or more time servers operating in this mode send periodic NTP broadcasts. All the workstations then determine the time on the basis of an assumed delay in the order of a few milliseconds. The advantage of these modes is that clients need not be configured for a specific server. Broadcast mode requires a broadcast server on the same subnet while multicast mode requires appropriate support for IP multicast. In this mode the server expresses its willingness to provide synchronization to many other peers, but does not accept NTP messages from any of them.

3.5 Data Formats

NTP carries out all arithmetic operations in two's complement fixed point format. Data are specified as integer or fixed point quantities. NTP timestamps, which represent the main product of the protocol, have been assigned a special format. When an event which has to be timestamped occurs, such as the arrival of a message, the current value of the local clock is converted to the above format and assigned to a timestamp variable. An NTP timestamp is a 64 bit unsigned fixed point number, with the integer part in the first 32 bits and the fractional part in the last 32 bits. The value is interpreted as the number of seconds, relative to UTC timescale, since 0H, January 1, 1900. The precision of this representation is about 200 picoseconds (2^{-32}) which should satisfy the requirements of every conceivable practical application to which NTP could be put to. Since 1968 the most significant bit of the 64 bit field has been set and the field will overflow sometime in 2036. If NTP is to be used after 2036, some external means would have to be put forth to distinguish time relative to 1900 and time relative to 2036. This format has been chosen as it facilitates multiple precision arithmetic and can be easily converted to other formats used by various protocols of the Internet suite, if necessary.

Fig 3.3 shows the NTP packet header format. Most of the fields are self explanatory. Leap indicator (LI) is a two bit code warning of an impending leap second to be inserted or deleted in the UTC timescale at the end of the current day. These are set before 23:59 H of the current day and reset after 00:00 H on

LI	VN	mode	stratum	poll	precision
Root delay - 32 bits					
Root dispersion - 32 bits					
Reference identifier - 32 bits					
Reference timestamp - 64 bits					
Originate timestamp - 64 bits					
Receive timestamp - 64 bits					
Transmit timestamp - 64 bits					
Authenticator (optional) - 96 bits					

Fig 3 3 NTP Packet Header

the following day Poll represents the interval between NTP messages sent. Each server uses the minimum of its own poll interval and that of the peer. This field contains a signed integer, in seconds as a power of two. The poll interval is adjusted dynamically to reflect dispersive delays and reachability failures. It varies from a minimum of 64 seconds (poll=6) to a maximum of 17 minutes (poll= 10). Root Dispersion indicates the maximum error relative to the primary reference source, while Reference clock identifier identifies the particular reference clock. Reference timestamp represents the time when the local clock was last updated. If the local clock has never been synchronized the value is zero. Authentication is an optional specification. If it is implemented the authenticator field contains the encrypted checksum of the message contents and the key identifier. The NTP authentication mechanism, which operates at the application level, is designed to protect against unauthorized message stream modification and misrepresentation of source. A more complete description of all the state variables used in the protocol machine can be found in the formal description[8].

3 6 NTP Functioning Model

Fig 3 4 shows the overall organization of the time server model. NTP

packets are exchanged between the host and other subnet peers. These are used to determine individual roundtrip delays and clock offsets relative to each peer. As shown in the figure, the computed delays and offsets are processed by the clock filter to cast out noisy data and the most accurate and reliable subset of peers determined by the clock selection algorithm. The resulting offsets of this subset are first combined using the clock combination techniques and then processed by a phase-lock loop (PLL). The local clock is modeled as a voltage controlled oscillator (VCO), which furnishes the timing reference to produce the timestamps used in all operations.

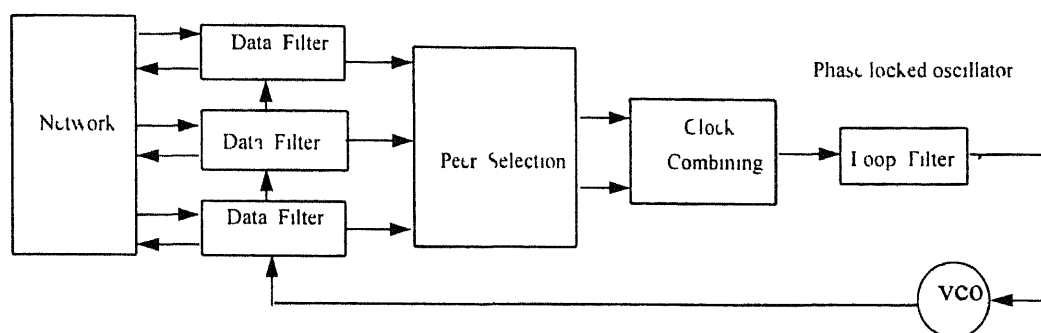


Fig 3.4 Time Server Model

3.6.1 Delay and Offset Computation

A peer timer is dedicated to each NTP association. When the peer timer decrements to zero (it is initially set with the appropriate “poll” value) the NTP client calls the transmit procedure and exchanges NTP packets with the peer. The timestamps thus exchanged enable the host to determine individual roundtrip delays and clock offsets, as well as to provide reliable error estimates. Clock offset is the time difference between the two clocks at a particular epoch, while the frequency difference between them is termed the skew. NTP is designed to produce three products: clock offset, roundtrip delay and dispersion. Since most host time servers will synchronize to the root via another peer time server, there are two components in each of these three products, those determined by the peer relative to the primary reference source of standard time.

and those measured by the host relative to the peer

Fig 3 5 shows how NTP timestamps are exchanged between host A and peer B. Let T_1, T_2, T_3, T_4 be the values of the four most recent timestamps as shown. With reference to NTP packet format of fig 3 3, T_1 corresponds to the originate timestamp, T_2 corresponds to the receive timestamp and T_3 to the transmit timestamp. T_4 is the time determined by host A upon arrival of the NTP message. Let

$$a = T_2 - T_1 \quad \text{and} \quad b = T_3 - T_4$$

The roundtrip delay δ is given as $[T_4 - T_1] - [T_3 - T_2]$
 $= a - b$

The roundtrip delay determined in this manner will be independent of the clock offset between the host and the peer. Under the assumption that the delay difference from A to B and from B to A, called differential delay, is small we further have

$$\theta = [T_3 + \delta/2] - T_4$$

$$= (a+b) / 2$$

where θ as shown in the figure is the clock offset of B relative to A at time T_4 . It can be shown that the true clock offset must lie in a window of size equal to the measured delay centered around the measured offset [8]

$$\text{i.e.} \quad \theta - \delta/2 \leq \theta_{\text{TRUE}} \leq \theta + \delta/2$$

Thus, both peers A and B can independently calculate delay and offset using a single bidirectional message stream. The advantages are that it does not depend on either the transmission times or the received message orders and

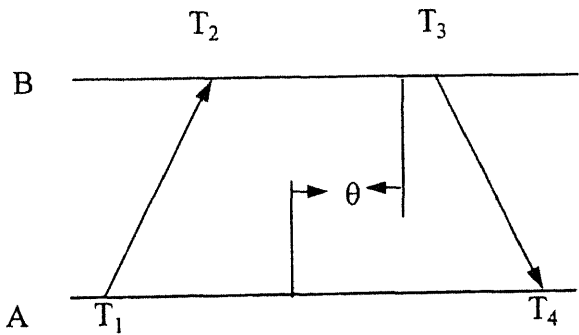


Fig 3 5 Measuring delay and offset

reliable delivery is not required. Also the dispersion, or the maximum error ε of the host relative to the peer is calculated as

$$\varepsilon = \rho + \varphi (T_4 - T_1)$$

The first term ρ gives an estimate of the measurement errors at the host. Each time the local clock is read a reading error is incurred due to the finite granularity or precision of the implementation. The second term $\varphi (T_4 - T_1)$, called the skew dispersion, gives a measure of the local clock skew accumulation over the interval since the last message was transmitted to the peer. φ represents the maximum skew rate constant. The dispersion thus calculated will be subsequently updated by the clock filtering procedure.

3.6.2 Data Filtering

The specification of the clock filter and clock selection algorithms do not constitute an integral part of the formal NTP specification since there could be many algorithms giving a satisfactory performance. The algorithms described here have been found to work well in the Internet environment. These have been devised after several years of experimentation with typical Internet paths. They have been engineered to provide high accuracy with low computational burden. The NTP data filtering algorithm, given a series of observations, attempts to produce an accurate estimate of the offset relative to a single peer. It belongs to the general class of convergence algorithms which attempt to reduce errors by repeatedly casting out statistical outliers. The clock filter procedure takes arguments of the form $(\theta_i, \delta_i, \varepsilon_i)$ where θ_i is a sample clock offset measurement and δ_i and ε_i are the associated roundtrip delay and dispersion. It subsequently determines the filtered clock offset, roundtrip delay and dispersion.

From various experimental results, it was concluded that the best offset samples occur at the lowest delays. This led to the design of a minimum filter which selects from the n most recent samples the sample with the lowest delay. It incorporates a shift register of NTP_SHIFT (an implementation defined constant, normally 8) stages, with each stage containing a 3-tuple $(\theta_i, \delta_i, \varepsilon_i)$. New data samples are shifted into the filter from the left resulting in old

samples falling of the right end All the NTP SHIFT samples are inserted in a temporary list and sorted in order of increasing delay (δ_i) The first sample in the list (θ_0 , δ_0 , ε_0) is presented as the filtered output The dispersion is however updated by taking into account a measure of the recent sample variance

The measure is called the filter dispersion (ε_σ) It is based on first order differences which are easy to compute, but at the same time forms a good quality indicator Consider the above sorted temporary list If the list has n entries (i e NTP SHIFT) in order of increasing δ_i , the filter dispersion ε_σ is defined as

$$\varepsilon_\sigma = \sum_{j=0}^{n-1} |\theta_j - \theta_0| v^j$$

where v is an experimentally adjusted weight factor The dispersion is subsequently updated

$$\varepsilon_0 \leftarrow \varepsilon_0 + \varepsilon_\sigma$$

3 6 3 Peer Selection and Combining Algorithms

The key factor for the high reliability associated with the NTP is the proper selection of the peer selection and combining algorithms When new offset estimates are produced for a peer or are revised as the result of timeout, this mechanism is used to determine which peer should be selected as the synchronization source and how to adjust the local clock, stratum and related variables

As noted earlier, the data filtering algorithm produces a triplet ($\theta_i, \delta_i, \varepsilon_i$) for the i th peer While these quantities are relative to the i th peer, the host also maintains a corresponding set of variables relative to the root of the synchronization subnet via the i th peer This set is denoted is (Θ_i, Δ_i, E_i)

These variables are computed as follows -

$$\Theta_i = \theta_i$$

$$\Delta_i = \text{peer rootdelay} + \delta_i$$

$$E_i = \text{peer rootdispersion} + \varepsilon_i + \varphi t$$

$$\Lambda_i = E_i + |\Delta_i|/2$$

The last two quantities E_i , Λ_i are referred to as synchronization dispersion and synchronization distance respectively φ as before is the maximum skew rate and t is the interval since last update. The clock selection process operates on the above quantities. Taking into account that the highest reliability is usually associated with the lowest stratum and synchronization dispersion while the highest accuracy is usually associated with the lowest stratum and synchronization distance, an NTP peer selection algorithm based on maximum likelihood statistical principles has been chosen.

The peer selection algorithm first checks for the validity of the data obtained from all the peers. If no peers pass the sanity check, the existing synchronization source, if any, is cancelled and the local clock free runs at its intrinsic frequency. After removing the defecting peers, if any, the resulting peers are sorted first by stratum and then by synchronization distance. Let $m > 0$ be the number of candidates in the sorted list. Θ_i is the offset associated with the i th peer. For each i ($0 \leq i < m$) the select dispersion ε_{ξ_i} relative to the i th peer is defined as

$$\varepsilon_{\xi_i} = \sum_{k=0}^{m-1} |\Theta_i - \Theta_k| \omega^k$$

where ω is an experimentally determined factor to cast out the outliers. The algorithm proceeds as follows. From the sorted list, discard the candidate with maximum ε_{ξ_i} , in case of ties discard the peer corresponding to maximum i .

Repeat the above procedure. The procedure terminates when the maximum select dispersion over all candidates left over in the list is less than the minimum filter dispersion of any candidate or until only a single candidate remains. The resulting effect is that the algorithm tries to select those peers at the beginning of the sorted list, which are at the lowest stratum and lowest delay and presumably can provide the most accurate time.

The result of the NTP clock selection procedure is a set of survivors (there will be at least one) all of whose times are statistically equal. Any of them can be chosen to adjust the local clock. However NTP specifies an optional clock-combining procedure, which has its roots in the one used by national standards laboratories to determine a synthetic laboratory timescale from an ensemble of cesium atomic clocks. In this procedure the offsets of the peers remaining on the list are combined with a weighted average algorithm to produce a more accurate working offset. The offsets of the peers are weighted by the estimated error, the error estimate being taken as the reciprocal of synchronization dispersion. It has been found experimentally that the clock combining procedure has been successful in considerably reducing errors in certain cases. A common problem in synchronization subnets is the offset errors resulting from asymmetric transmission paths, where the network paths in one direction are substantially different from the other. These kind of errors have been reduced by implementing the above optional specification.

The result of the clock selection and clock combining procedures is the final clock correction Θ , which is used by the local clock procedure to update the local clock. It also sets the local stratum to one greater than the stratum of the selected peer. The synchronization distance and dispersion are all calculated and stored as system variables. Using the previous notation and assuming the i th peer is selected for synchronization, the system variables are determined as follows

$$\Theta = \text{combined final offset}$$

$$\Delta = \Delta_i \quad \text{and} \quad \Lambda = \Lambda_i$$

$$E = E_i + \delta_{\xi_i} + |\Theta|$$

3.7 The NTP Local Clock

In order to attain the full accuracy and reliability that the NTP can support, careful considerations must be given to the design of local clock hardware and software. In the NTP model the local clock is assumed to function with an uncompensated crystal oscillator. However the timekeeping performance can be substantially increased by using a precise and stable local oscillator. The NTP logical clock model is shown in fig 3.6. The difference between the peer time and the server time $T_B - T_A$ i.e. the offset θ determined from the exchange of timestamps is processed by the phase detector to produce the output V_D . The filtering, selection and combining algorithms which maintain a set of previous offsets are modeled as a variable delay network to produce the output V_S . The loop filter produces the output voltage V_C to control the VCO so as to reduce the offset. The above model is viewed as a type II adaptive parameter phase lock loop (PLL) which continuously corrects the local oscillator phase and frequency in order to compensate for its intrinsic jitter, wander and drift, from the offsets determined by the selection and combining algorithms.

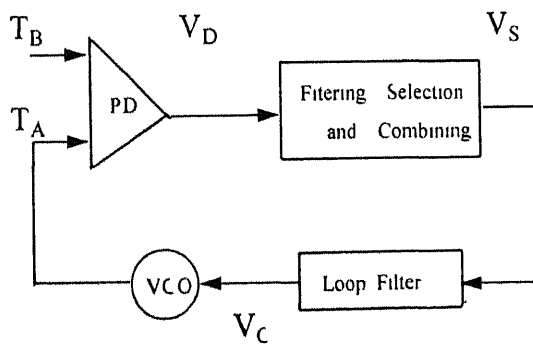


Fig 3.6 NTP Clock Model

The open loop transfer function of a type II PLL is taken as

$$G(s) = \left[1 + \tau s / \omega_z \right] \omega_c^2 / \tau^2 s^2$$

where ω_c is the crossover frequency, ω_z is the corner frequency and determines

the loop stability and τ is the PLL time constant. Under the assumption that the induced delay by the data filtering, selection and combining network is small, the loop filter transfer function can be written as [8]

$$F(s) = 1/K_s \tau + 1/K_f \tau^2 s$$

$$K_f = \alpha / \omega_c^2 \quad \text{and} \quad K_s = K_f \omega_z$$

Here α is the VCO gain. In the implementation model, the reciprocal of α is taken as the adjustment interval, i.e., the time between two clock corrections. The τ parameter determines the PLL time constant and thus the loop bandwidth. Using the α and τ parameters, the PLL adapts its behaviour to match the prevailing stability of the local oscillator and the transmission conditions of the network.

The method in which the local clock and other parameters are updated can be seen in the following analysis based on a set of recurrence equations. The PLL time constant τ is computed from a parameter called compliance (h). The concept of compliance is similar to that of Allan variance used in time-keeping systems based on cesium clocks. The compliance is a measure of prevailing time dispersions.

The local clock is continuously adjusted in small increments at fixed intervals of σ . On receipt of NTP messages, the update procedures are invoked and the various state variables are updated. The update intervals μ are variable and can vary up to 17 minutes. Fig. 3.7 shows a section of the timescale when the i th update has been received. The updates are numbered from zero, and all the variables are suitably initialized at $i=0$. From the figure, the i th update

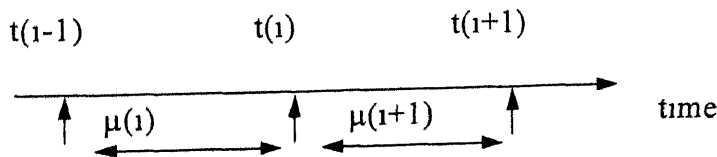


Fig. 3.7 Update Nomenclature

arrives at $t(i)$ and has an associated time offset $V_s(i)$ (fig 3.6). The $(i+1)$ th update arrives at $t(i+1)$ after an interval of $\mu(i+1)$, and $V_s(i+1)$ is its associated time offset. When the updates $V_s(i)$ is received the frequency error $f(i+1)$ and phase error $g(i+1)$ are computed as

$$f(i+1) = f(i) + \mu(i) V_s(i) / \tau^2(i)$$

$$g(i+1) = V_s(i) / \tau(i)$$

The time constant to be used for the next update is computed from the current value of compliance $h(i)$

$$\tau(i+1) = \tau^l = \text{Max}[K_s - |h(i)|, 1]$$

The updated poll interval is set to $K_u \tau^l$. The value of compliance to be used in the $(i+1)$ th update is arrived according to

$$h(i+1) = h(i) + [K_l \tau^l V_s(i) - h(i)] / K_h$$

The appropriate values of the parameters K_s , K_u , K_h , K_l , in what way the state variables have to be initialized, and other details can be found in [8]

As noted above the factor τ^l controls the PLL bandwidth. When the compliance has been low for a long period, τ^l is increased and the bandwidth is decreased. In this state small timing fluctuations due to jitter in the network are suppressed and the PLL attains the most accurate phase estimate. On the other hand if the compliance becomes high, say due to network congestion or a systematic frequency offset τ^l is decreased resulting in increased loop bandwidth. In this mode the PLL adapts itself quickly to these transients. The NTP daemon, which is considered in the succeeding chapter, simulates the above recurrence relations and provides offsets to the kernel at intervals of $\sigma = 1$ second using the `adjtime()` and `ntp_adjtime()` system calls.

3 8 Present Status

In the last few years the NTP has become the most widely deployed protocol in the Internet for accurate timekeeping. The NTP performance issues are discussed in [11]. A newer version of NTP, version 4, has been proposed [12]. The proposed new version contains several improvements to existing algorithms apart from some new features to support multicast modes of operation, fully distributed peer subnets and automated NTP peer discovery and subnet configuration. It is yet to become an Internet standard.

Chapter 4

GPS Time and Position Outputs: Studies and some Applications

In this chapter we describe the work that we have carried out. The first section shows how we have benefitted from the GPS time in improving the Network timing accuracy. The next section dwells upon some of the studies that we have carried out on the GPS receiver outputs.

4.1 Implementation of a Time Server

We have set up a stratum-1 NTP Primary time server synchronized to a GPS clock. The Accords GPS Receiver GPSR 2000 was employed as a source for the GPS timing signals. We have used the Xntp3-5.90 distribution for setting up the time server which is available from [5]. The Xntpd daemon is a complete implementation of the Network time protocol, version 3 specification. The xntpd daemon was appropriately interfaced and configured to support the Accords GPS Receiver. The Time server software was run on a Linux platform. First we look at the mechanism through which NTP supports these Primary time standards and then consider the overall implementation details.

4.1.1 Reference Clock Support

The NTP daemon, Xntpd, has been designed to support external clocks such as those derived from radio, satellite or modem clocks. These clocks are handled by the protocol in a manner similar to that of ordinary NTP peers. A peer is uniquely characterized in a network through its 32 bit IP address and the

NTP manages all its associations based on the peer's IP address. In a similar way, the host clocks are also assigned IP addresses. These are syntactically correct, but are invalid IP addresses, so chosen so as to distinguish them from the normal NTP peers. The clock addresses are of the form 127.127.t.u, where t specifies the particular type of clock, i.e., refers to a particular clock driver and u is a number in the range 0-3, intended for supporting multiple instances of clocks of the same type on the same server. Thus, should redundancy be required, NTP supports it, perhaps much more than required. In essence the reference clock support is provided by maintaining the fiction that the clock is actually a peer, but is assumed to operate at a synthetic stratum of zero. As no packets are exchanged with a reference clock, the transmit, receive and packet procedures corresponding to a normal peer association are replaced by a separate code to simulate them i.e. this code maintains the state variables corresponding to the reference clock association in an analogous way as that of any other peer. So the entire suite of algorithms used to filter the received data, select the best clocks or peers and combine them to produce a local clock correction are operative as usual. In this way defective clocks can be detected and removed from the peer population.

4.1.2 Running The NTP Daemon

The daemon can be started and stopped from root anytime. However the command line that starts the daemon is normally included in the system startup file (/etc/rc.d/rc.local), so it is executed at the system boot time. On being initiated the daemon first looks for a configuration file (/etc/ntp.conf) to find out the servers or peers that it can poll in order to obtain the timing information. The configuration file is an ASCII file conforming to the usual comment and whitespace conventions. We employed the following configuration file for running the daemon.

```
# XNTPD Configuration file
#gps server
server 127.127.20.1
```

```
#Two departmental servers  yamuna and chambal running the NTP
server 202 141 40 3
server 202 141 40 4
#driftfile declaration
driftfile /etc/ntp drift
```

The GPS clock was configured with the address shown. Two other systems in the network were running NTP configured to obtain the time from one of the time servers in the Internet. Since the host is expected to operate at stratum-1 these two servers should have at the best been declared as peers, but the above one was chosen just to test the software. The configuration file also includes a driftfile declaration. On being initiated, NTP first tries to compute the error in the intrinsic frequency of the clock on the host it is running on. It needs a good estimate of this to synchronize closely to the selected servers and to converge on to a good estimate of the frequency error, it usually takes a day or so after the daemon is started. Once the initial value is arrived at, it will change only by relatively small amounts during the course of continued operation. The driftfile declaration indicates the name of the file where the NTP may store the current estimate of the computed frequency error, so that, in case the daemon is stopped and restarted for some reason, it can reinitialize itself to the previous refined estimate and doesn't have to compute it all over again.

Using the above configuration file, the daemon was run on the host. The correct operation of the daemon was verified by using some of the monitoring facilities included in the distribution. Fig 4.1 shows all the associations the host maintains. All those configured as servers in the startup file show up in the remote column. The NTP polls all the servers listed and the one shown with the '*' is the currently selected server for synchronization. The refid entry shows the current source of synchronization (root) for each peer, while the st reveals its stratum and the poll entry, the polling interval in seconds. As explained earlier, the GPS clock is shown to operate at a stratum of zero. The other two peers are shown to operate at a stratum of 16, i.e., they are currently out of sync. The when entry shows the last time the peer was heard, in seconds. The reach

remote	refid	st	when	poll	reach	delay	offset	dispersion
yamuna	128 118 25 3	16	903	1024	377	4 21	18172	48 52
chambal	0 0 0 0	16	299	1024	0	0 00	0 00	16000 0
*GPS_ACCORD	GPS	0	14	64	377	0 00	110 79	6 20

Fig 4 1 Associations status

```

status=04f4 leap_none sync_uhf_clock 15 events event_peer/strat_chg
system='Linux' leap=00, stratum=1 rootdelay=0 00
rootdispersion=105 74, peer=50854 refid=GPS
reftime=b88d4853 1ad63000 Thu Feb 12 1998 15 37 15 104 poll=6
clock=b88d4853 57e3b000 Thu Feb 12 1998 15 37 15 343 phase= 104 831
freq= -77856 69 error= 1 07

```

Fig 4 2 Clock Status

entry shows the status of the reachability register (8 bits) in octal format. A peer is considered reachable if atleast one bit of the register is set to 1, thus the second server listed is unreachable. The latest samples of offset, delay and dispersion computed for each peer in milliseconds are also exhibited.

Fig 4 2 exhibits the details of the local clock. It shows the host as operating at stratum-1 synchronized to the GPS clock. The peer entry corresponds to the refclock association id. The clock entry corresponds to the time when this set of readings were taken, while the reftime corresponds to the time (the time is given in both the NTP timestamp format and in the local time) when the local clock was last updated. The daemon continuously tracks the discrepancy between the local time and the NTP time and adjusts the local clock accordingly. This adjustment occurs in two dimensions- time and frequency.

Fig 4 3 shows a clip of the messages logged into the system log. It shows the times when the NTP has stepped the local clock to reduce the error. Normally the daemon will adjust the local clock in small steps in such a way

```

12 Feb 09 19 36 xntpd[4969] logging to file syslog
12 Feb 09 19 36 xntpd[4969] xntpd 3 5 90 Thu Feb 12 09 17 18 IST 1998 (1)
12 Feb 09 19 36 xntpd[4969] tickadj = 1 tick = 10000 tvu_maxslew = 99, est hz = 100
12 Feb 09 19 36 xntpd[4969] precision = 17 usec
12 Feb 09 19 36 xntpd[4969] read drift of -72 587 from /etc/ntp drift
12 Feb 09 20 06 xntpd[4969] synchronized to GPS_ACCORD(1) stratum=0
12 Feb 09 20 04 xntpd[4969] time reset (step) -1 493148 s
12 Feb 09 20 04 xntpd[4969] synchronisation lost
12 Feb 09 20 11 xntpd[4969] synchronized to GPS_ACCORD(1) stratum=0
12 Feb 09 59 15 xntpd[4969] time reset (step) -0 214383 s
12 Feb 09 59 15 xntpd[4969] synchronisation lost
12 Feb 10 00 05 xntpd[4969] synchronized to GPS_ACCORD(1) stratum=0
12 Feb 10 37 00 xntpd[4969] time reset (step) 0 211654 s
12 Feb 10 37 00 xntpd[4969] synchronisation lost
12 Feb 10 37 05 xntpd[4969] synchronized to GPS_ACCORD(1) stratum=0
12 Feb 10 39 08 xntpd[4969] recvfrom() fd=7 Connection refused
12 Feb 11 16 03 xntpd[4969] time reset (step) 0 202369 s
12 Feb 11 16 03 xntpd[4969] synchronisation lost
12 Feb 11 16 08 xntpd[4969] synchronized to GPS_ACCORD(1) stratum=0
12 Feb 12 07 13 xntpd[4969] time reset (step) 0 195442 s
12 Feb 12 07 13 xntpd[4969] synchronisation lost
12 Feb 12 08 05 xntpd[4969] synchronized to GPS_ACCORD(1), stratum=0
12 Feb 13 26 02 xntpd[4969] time reset (step) 0 182863 s
12 Feb 13 26 02 xntpd[4969] synchronisation lost
12 Feb 13 26 07 xntpd[4969] synchronized to GPS_ACCORD(1) stratum=0

```

Fig 4 3 A Section of the Messages Logged

that system and user programs are unaware of its operation. The adjustment process operates continuously as long as the apparent clock error exceeds 128 milliseconds. If the apparent time error persists for an interval of about 20 minutes, the local clock is stepped to the new value. When the clock is reset, the clear procedure is invoked to clear all the clock filters, reset the poll interval

and reselect the synchronization source, if necessary. This explains the "synchronization lost" messages in the logfile after the clock is reset. The logfile also shows the initialization messages when the daemon is invoked.

4.2 GPS Receiver Measurements: A study

We have studied some of the characteristics of a commercial GPS Receiver outputs. In our setup we had used two Accord's GPSR 2000 Receivers. The two receivers were placed at the two farthest points (say locations '1' and '2') of our roof top, about 70 meters apart. A software driver was written to monitor the Accord's receiver binary outputs. This driver was used to log the computed outputs, and the data was processed offline. The first sub-section discusses the DOP parameters and the next one considers the positional output variances in standalone and differential modes.

4.2.1 DOP Measurements

As described in section 2.8, the relationship between the errors in the pseudorange measurements of the user and the user's position and time accuracy is described by the GDOP (Geometric dilution of precision). In this we considered the positional dilution of precision (PDOP) and the horizontal 1 e two dimensional component of it (HDOP). The DOP parameters are functions of the user/satellite relative geometry. Since the GPS satellites are in 12-hr sidereal time orbits the user/satellite geometry approximately repeats every 24 hours. Thus a 24-hr plot, more or less, characterizes the DOP at a particular location. We obtained a 24-hr data of the PDOP and HDOP parameters using two different receivers at the same location on two different days. Their plots (the DOP parameters were averaged over a window of 1 minute) are shown in fig 4.4 - 4.7. As expected the two sets of data do exhibit a degree of similarity.

In the next case the two receivers were placed at the two different locations 1 and 2. At each location, during the same period of time, the data was taken for about seven hours and the resulting plots are shown in fig 4.8 - 4.11. These plots as well as the previous ones considered exhibit a few

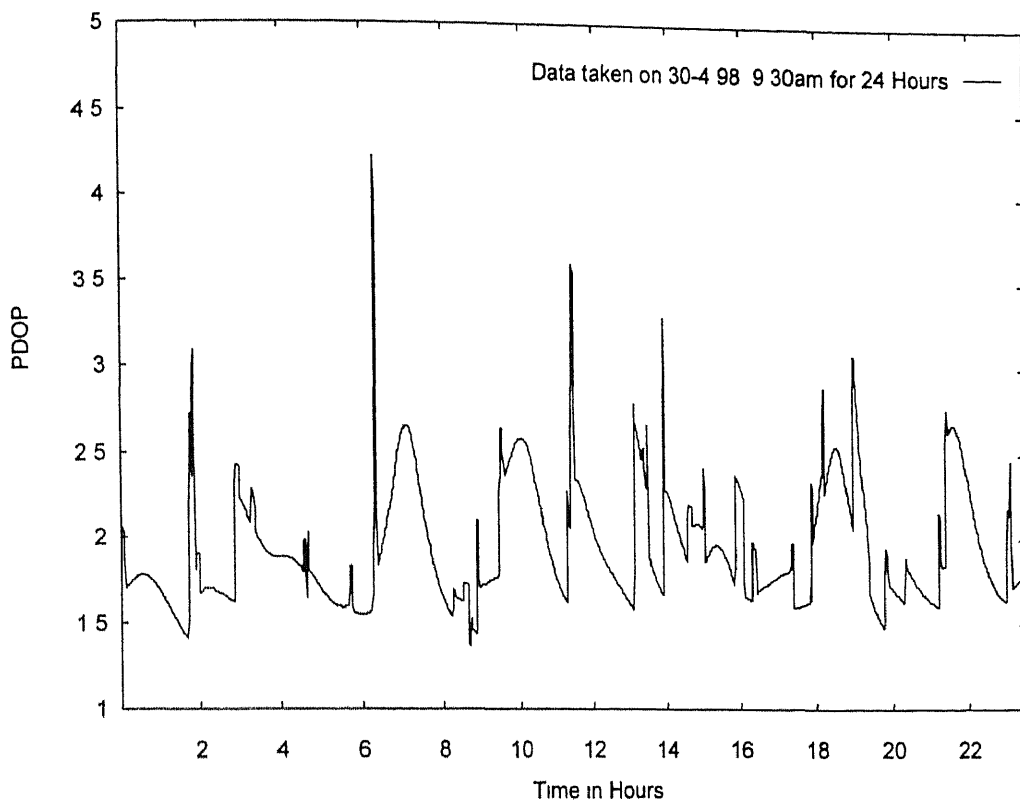


Fig 4 4 PDOP Variation obtained with Receiver 1

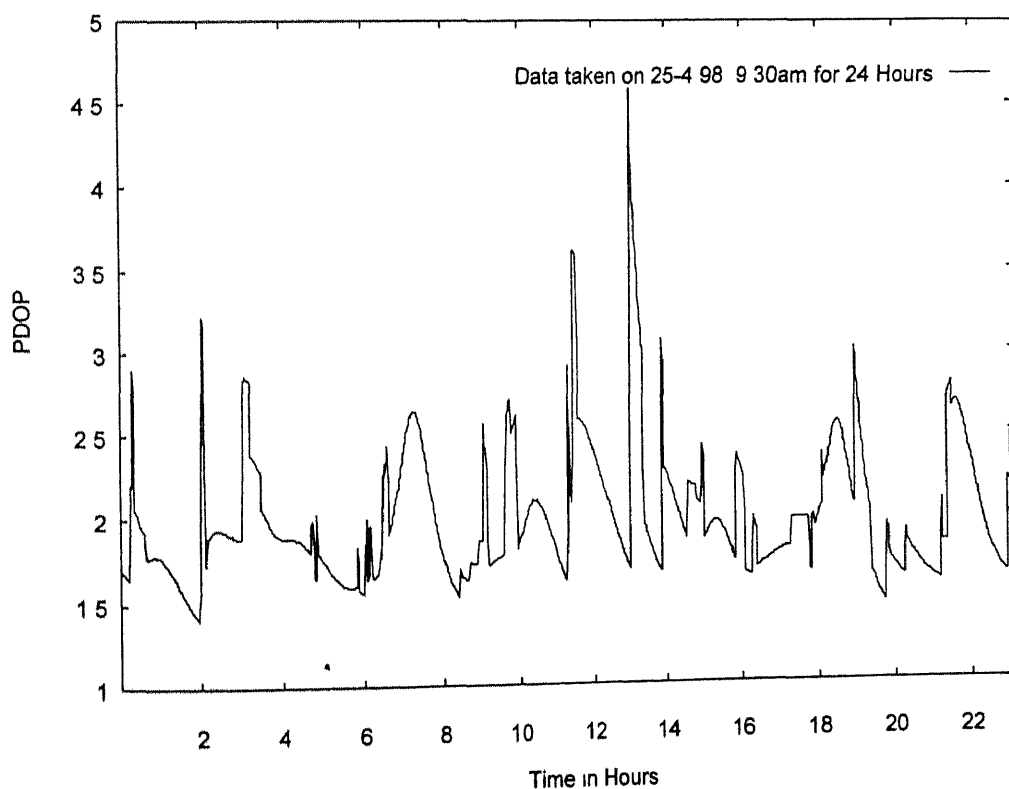


Fig 4 5 PDOP Variation obtained with Receiver 2

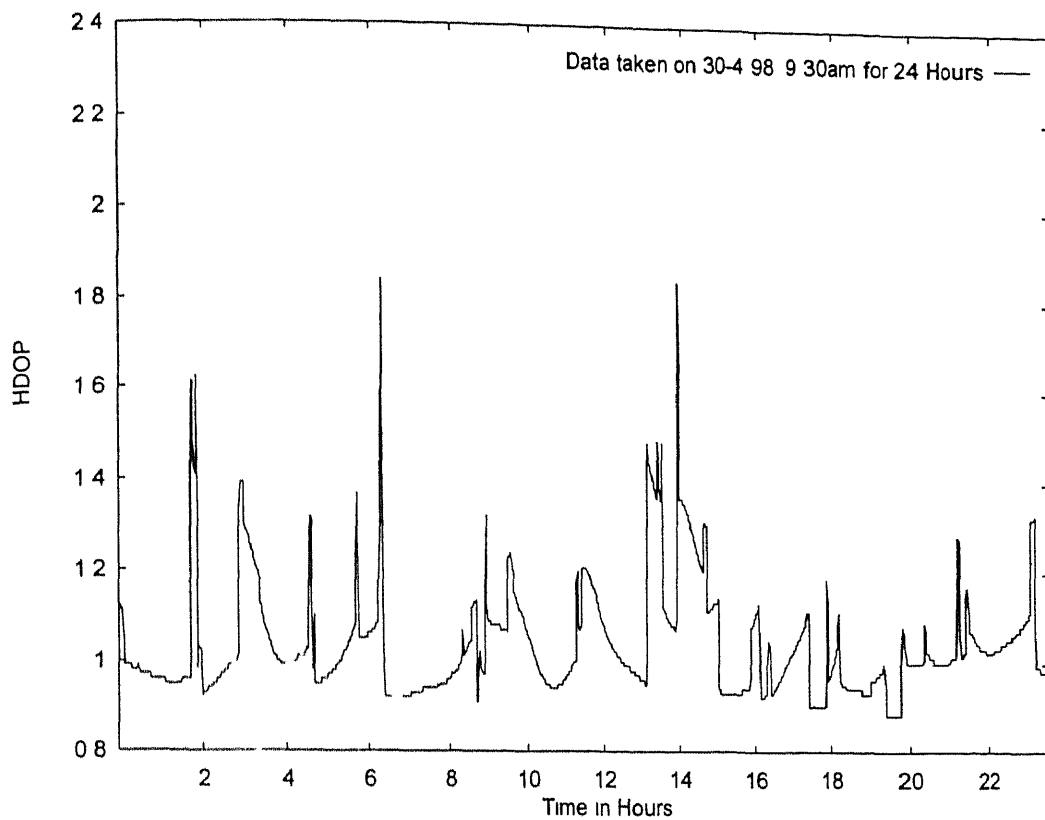


Fig 4 6 HDOP Variation obtained with Receiver 1

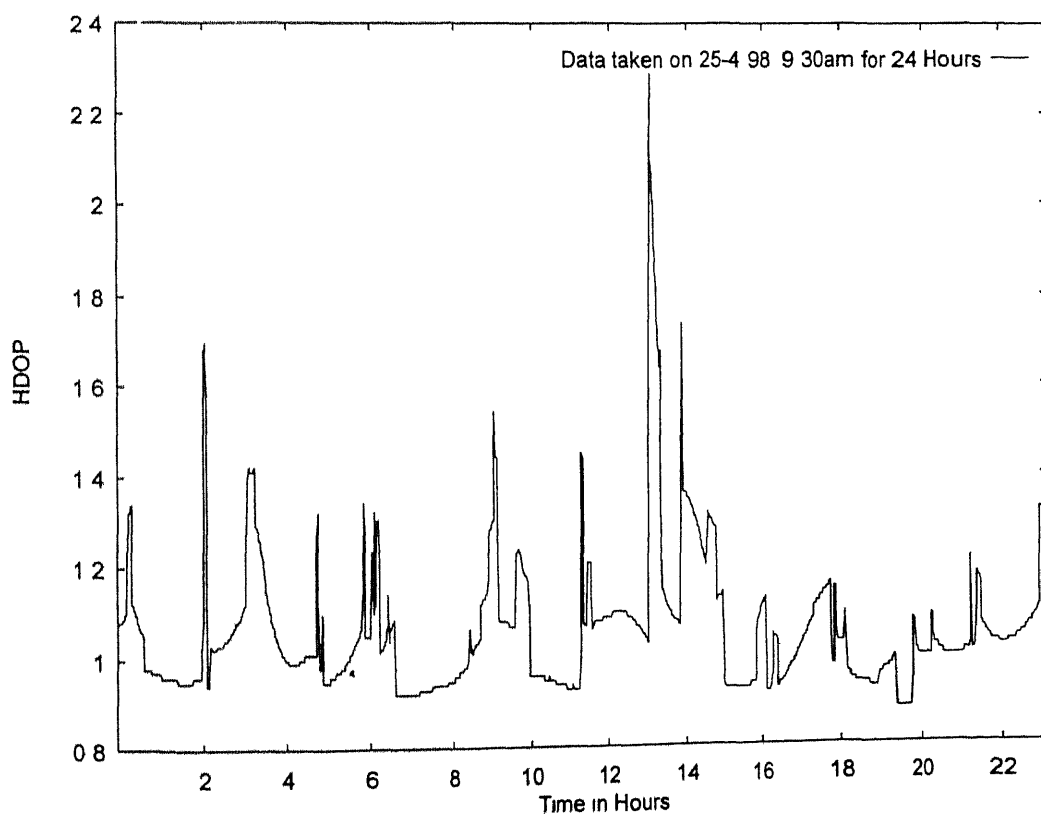


Fig 4 7 HDOP Variation obtained with Receiver 2

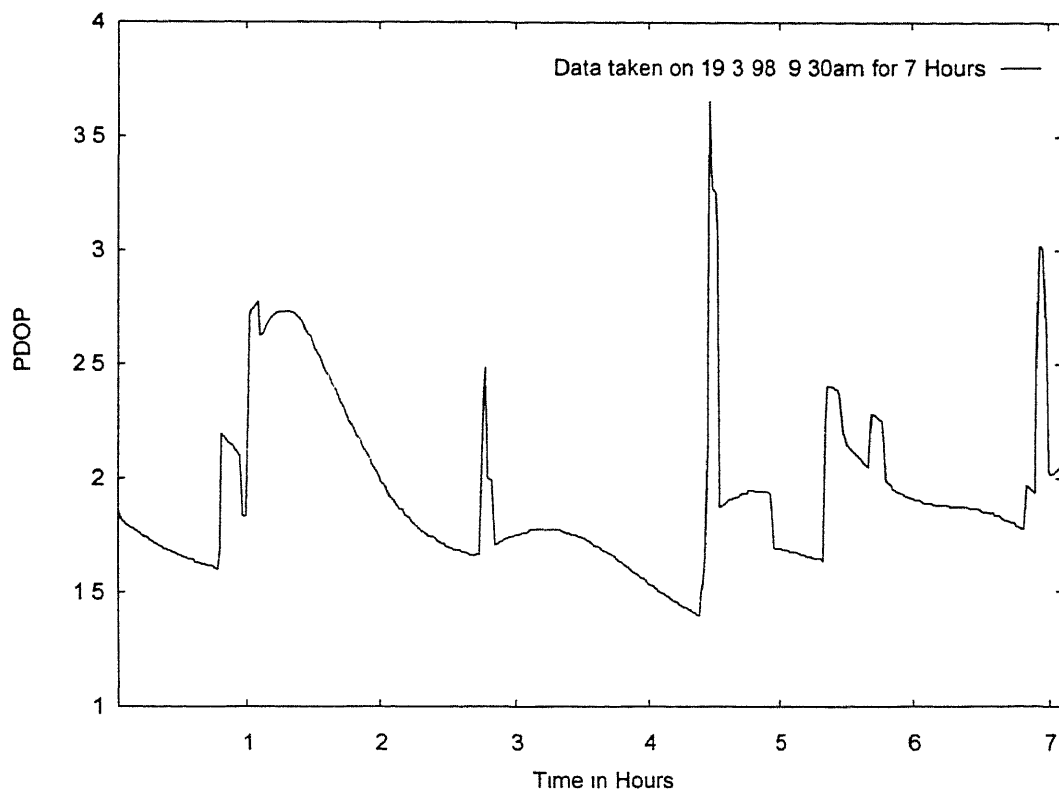


Fig 4 8 PDOP Variation at Location 1

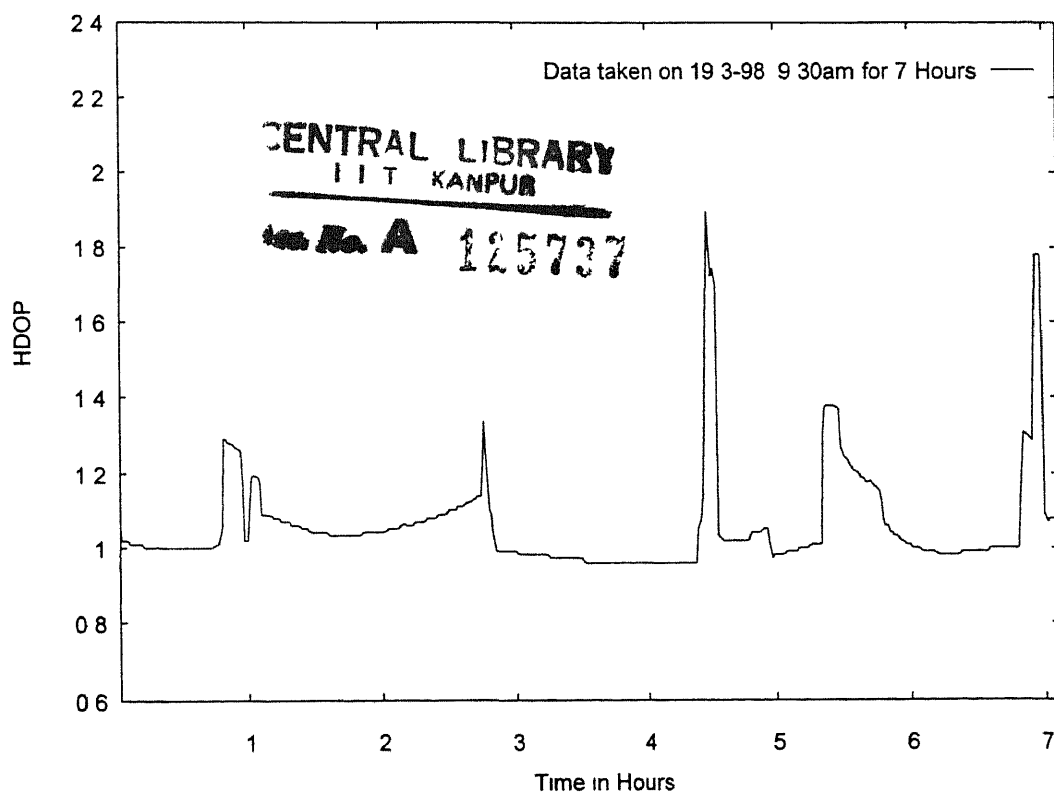


Fig 4 9 HDOP Variation at Location 1

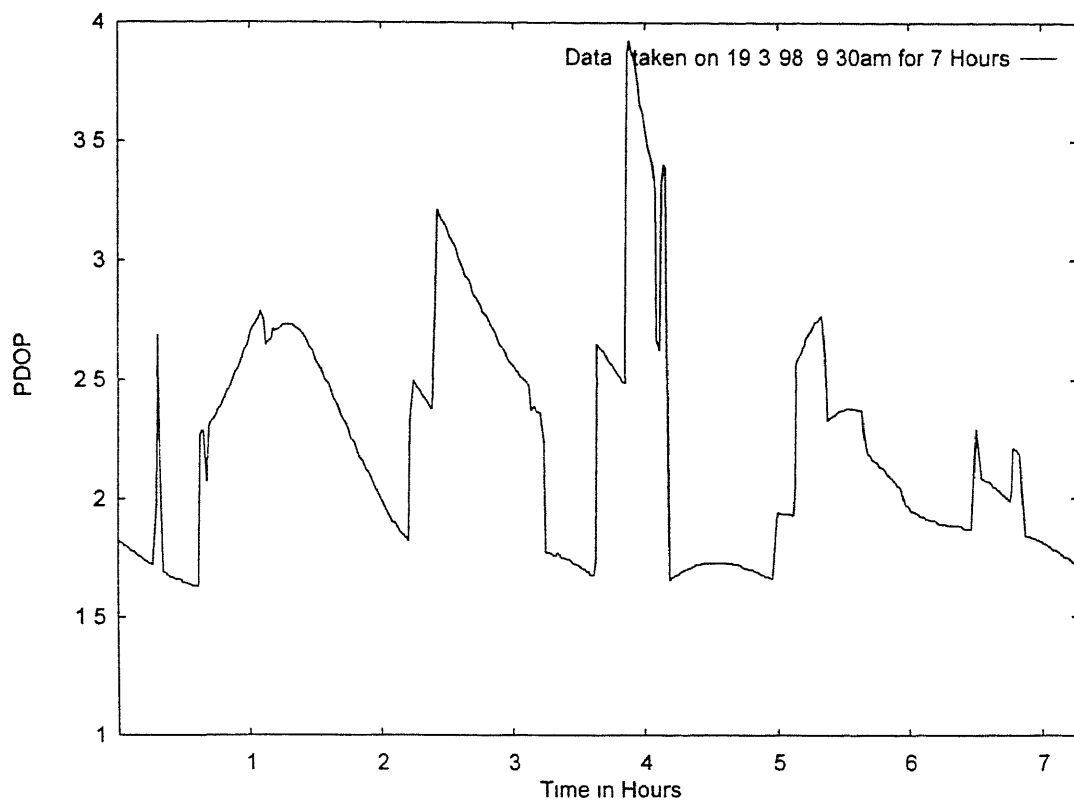


Fig 3 10 PDOP Variation at Location 2

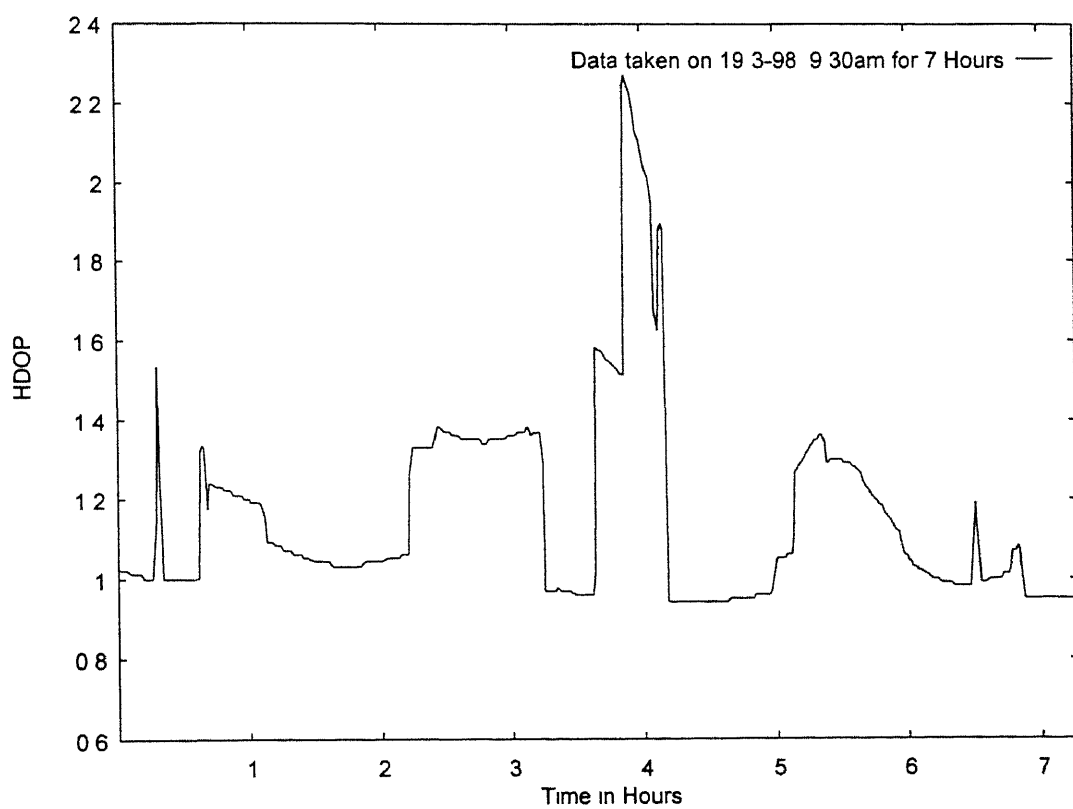


Fig 3 11 HDOP Variation at Location 2

discontinuities. The discontinuities in the plots possibly occur when the satellite configurations employed by the receiver in the navigation and time solution are changed. The number of satellites in view from a given location depends on the receiver mask angle ($0-10^\circ$) and varies as a function of time [3]. If a satellite currently in view and used by the receiver in the position solution moves below the mask angle, the receiver is forced to select another satellite with possibly a worse (higher) GDOP. On the other hand, if a satellite initially out of view moves above the mask angle, the receiver may choose to switch to it. Both these cases of satellite changeover result in an abrupt change in the DOP parameters.

4.2.2 Standalone and Differential Position Measurements

The goal of our doing these measurements was to get an idea of the differential GPS. As already explained, differential GPS techniques are used to tackle the inaccuracies resulting from the spatially correlated errors like the SA or the satellite clock errors. The best approach to compute differential GPS measurements is to let the comparison be done at the pseudoranges, from both the receivers, with reference to an identical set of satellites, using which the position measurements are calculated. Since the data at pseudorange levels referenced to satellite identities was not available from the commercial GPS receivers, an attempt to study differential measurements has been carried out using the processed position data available from the two GPS receivers. Fig. 4.12 - 4.17 show the plots generated taking the receiver's latitude output as a specific case.

The receiver recomputes its position estimate every second. As shown, a 7-hr data was considered and the statistical mean and variances were computed over a window of every 1 minute. Fig. 4.12 - 4.15 show the statistics computed separately for each location while Fig. 4.16 and 4.17 consider the differential modes. We had expected that some of the errors and biases common to locations 1 and 2 would get suitably negated and hence the differential variances to be considerably lower than the individual ones. However, since this analysis is only an approximation to differential techniques, the observed difference in variance was not all that marked. Still, in some regions of the graph, the differential variances are lower than the ones in the standalone mode.

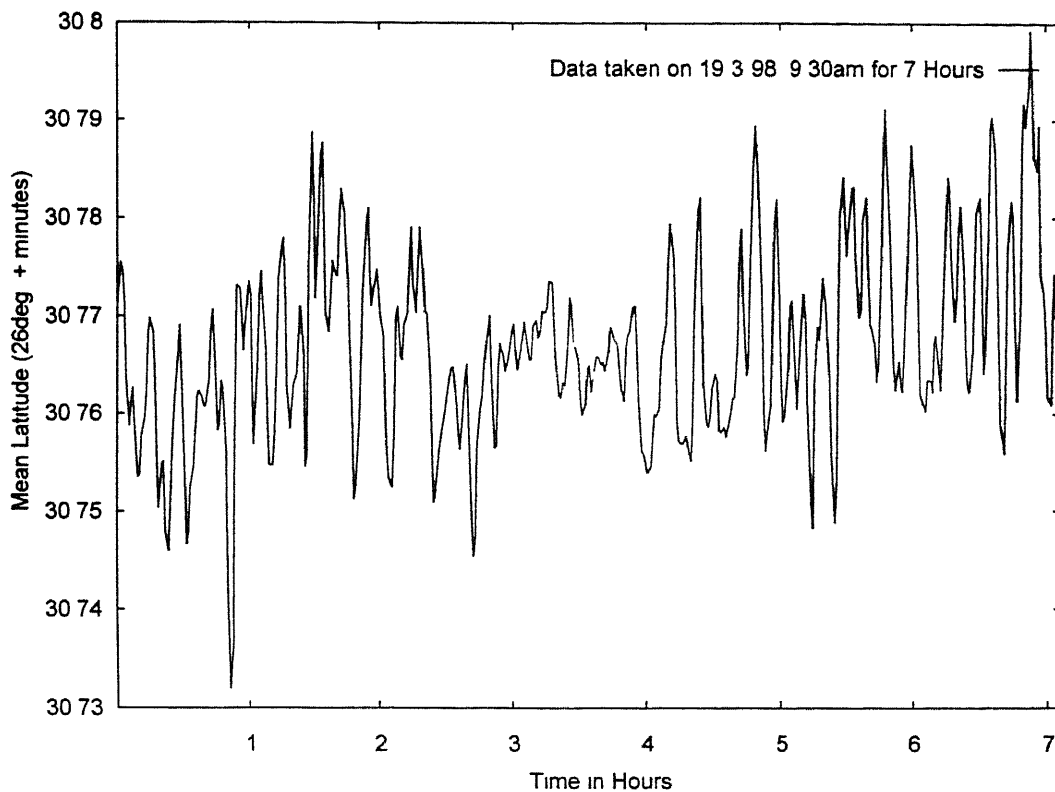


Fig 4 12 GPS Latitude output at Location 1

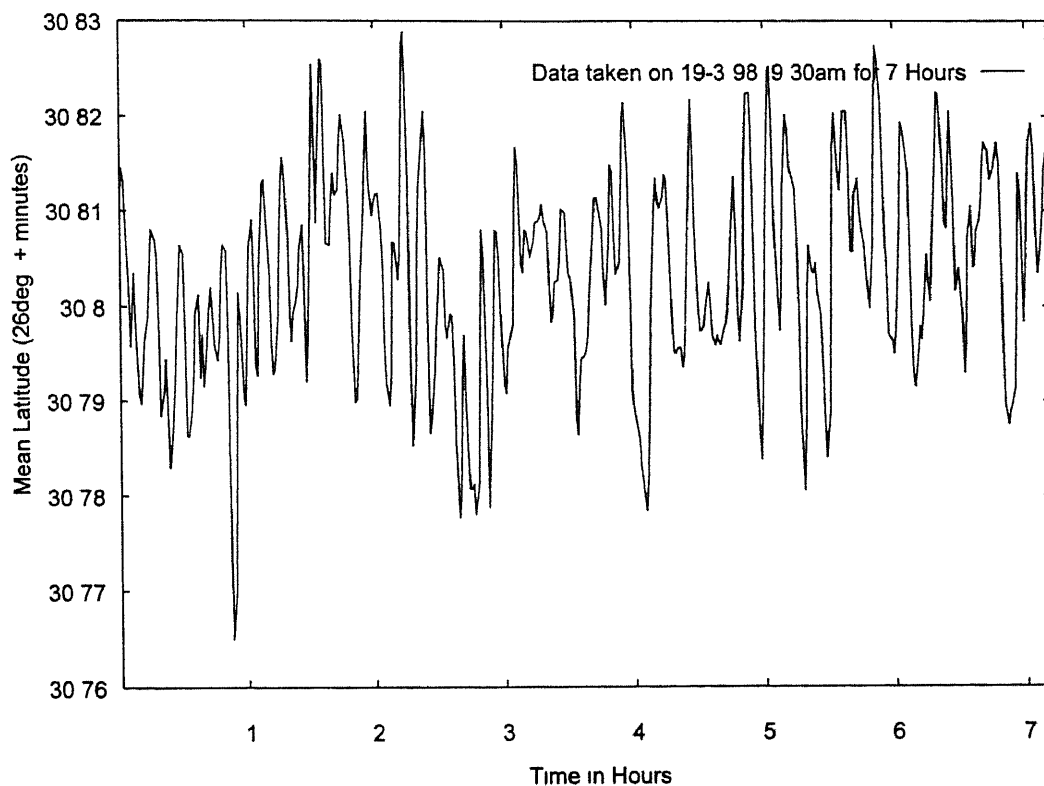


Fig 4 13 GPS Latitude output at Location 2

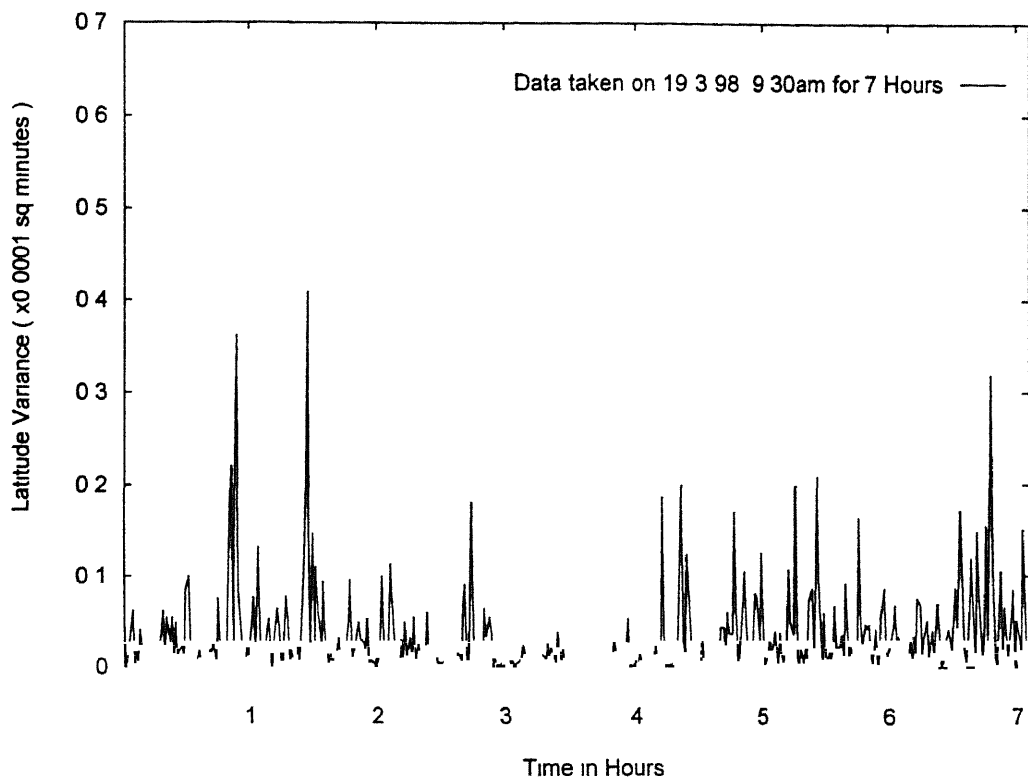


Fig 4 14 Latitude variance at Location 1

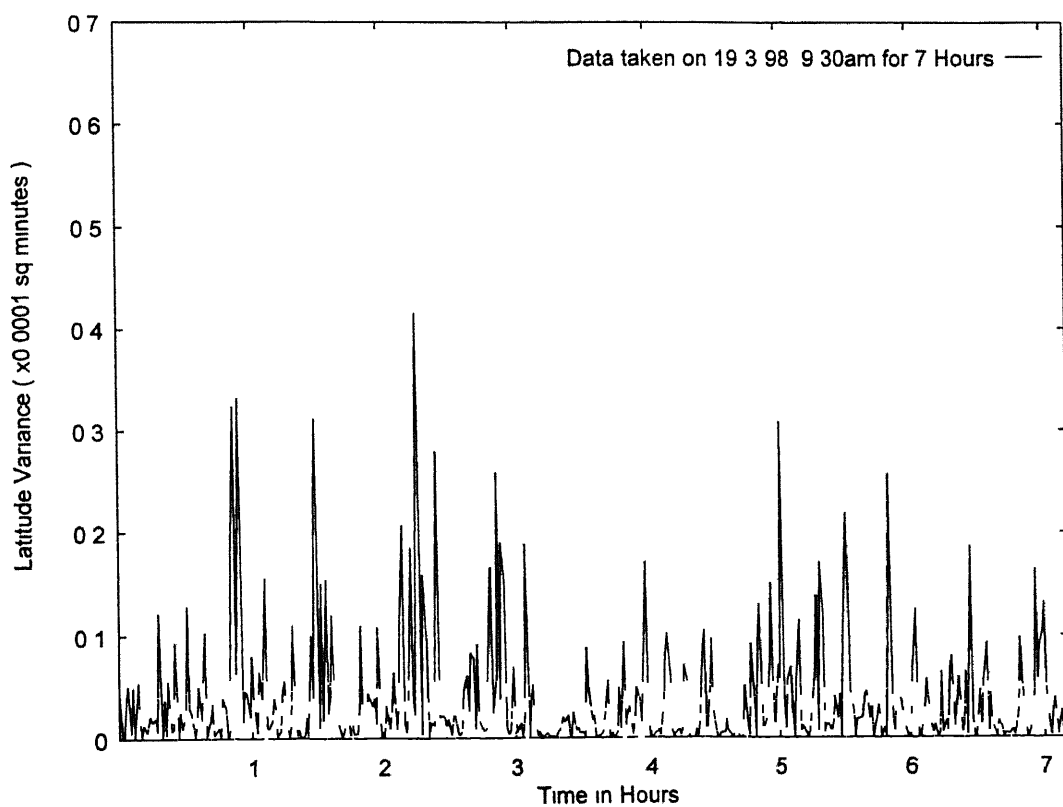


Fig 4 15 Latitude variance at Location 2

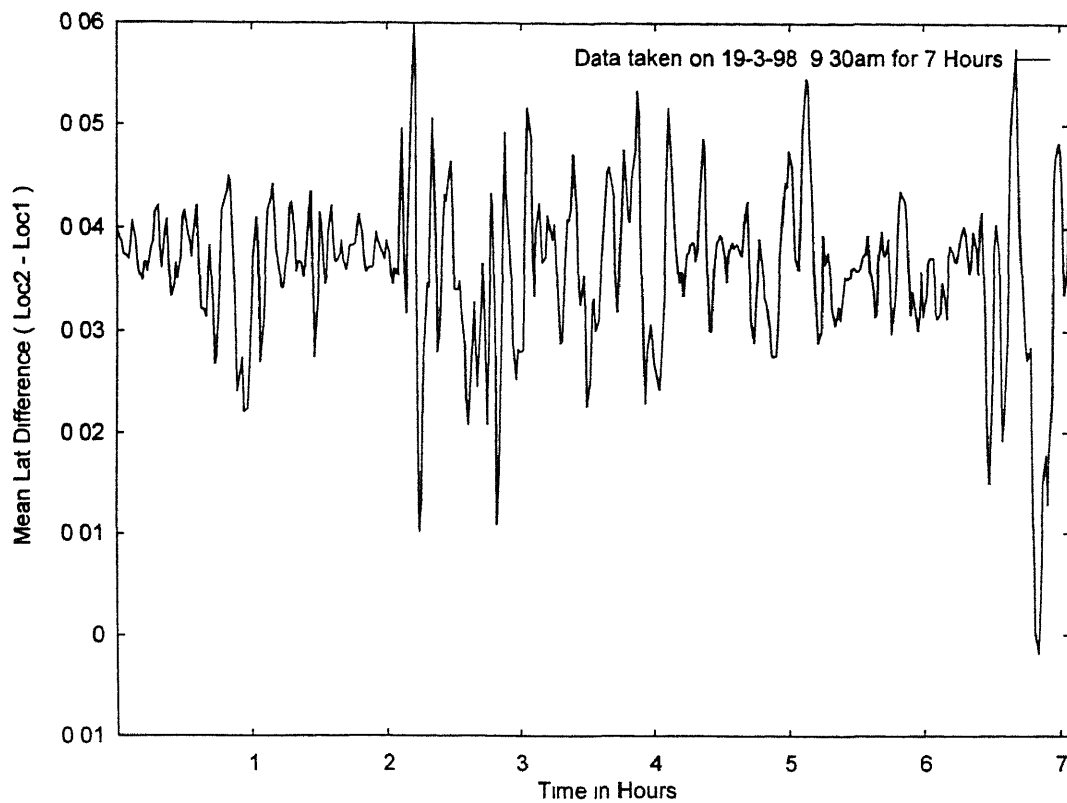


Fig 4 16 Differential Latitude Variations

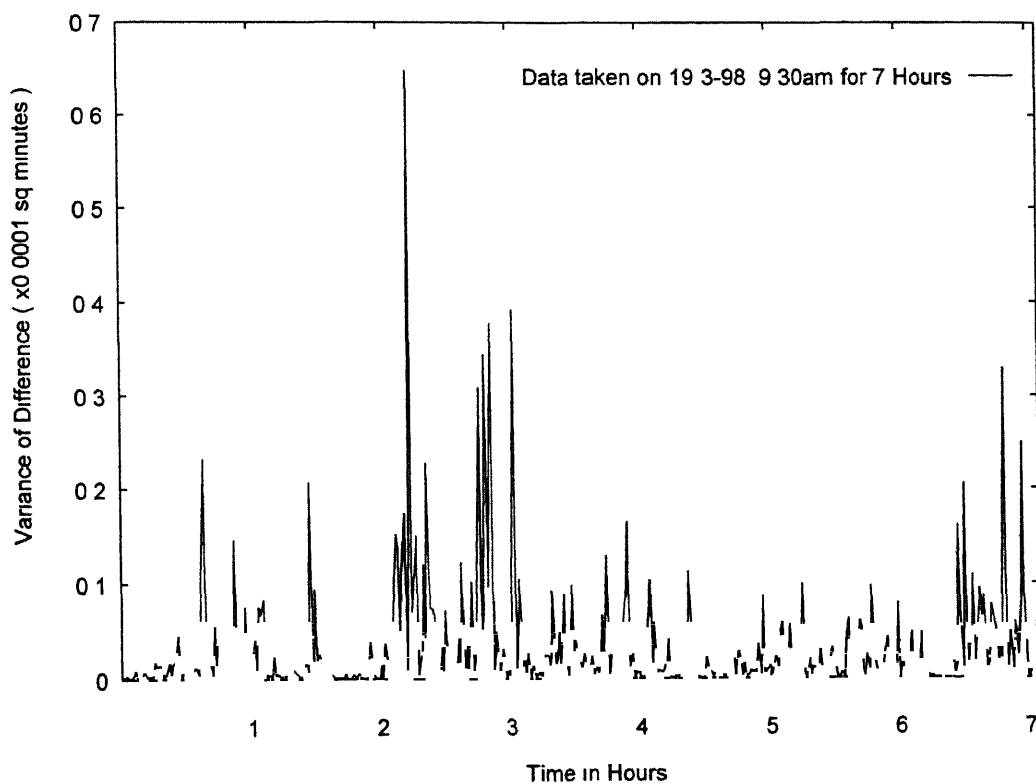


Fig 4 17 Latitude Variance in differential mode

Chapter 5

Conclusions

In this thesis work we have considered some of the ways in which GPS can be put to use in the networking area. We have shown that the GPS can be employed as an excellent reference for the Network timing. There is a growing set of computer applications that require precise synchronization to function properly. Before the advent of the Global Positioning System, it was financially impractical for most users to operate their own stratum one servers and, therefore, most users obtained their time from some remote time servers via one of the LAN protocols. In certain cases this may considerably affect the accuracy to which time can be maintained in the local network. During our observations we found that the two gateways of our local network, which were configured to obtain their time, via NTP, from some Internet time servers were out of synchronization for a considerable period of time in a day. We have used the GPS derived time to create a Primary Time server, running the Network Time protocol. This time server can now be used to co-ordinate the time distribution within the local network and outside. The stratum-1 server can maintain time within a few tens of milliseconds with reference to the world standard, the Universal time coordinated (UTC). It can be used to significantly increase the accuracy to which time can be maintained in the local network and thus increase the overall reliability of the network.

In this thesis we have also looked at the various outputs of a commercial GPS receiver. We have studied the characteristics of the DOP parameters and how they vary with time. We have also looked at the positional output variances in standalone and differential modes.

5 1 Suggestions for Further work

A newly emerging area like GPS offers ample opportunities for further work. The field is vast and open. A very large number of applications are cited in the GPS literature. Its presence is being increasingly felt in many areas including defence, aviation and various commercial sectors. Serious developments of strategic and commercial GPS applications will depend on several factors such as

- The investment we make in both ground and space segments in the form of local area and wide area augmentation systems
- Combined GPS and GLONASS Receivers
- Development and deploying of geodetic quality receivers
- Development of networked GPS/GIS applications
- Efforts to bring in wide awareness in the country through education and training and incentives to industry and social sector services to deploy GPS/GIS applications

Appendix A

This Appendix summarizes the Accords GPS Receiver output formats which have been used in this work. A complete description of the receiver's outputs can be found in the Receiver manual [4]. The processed data is output in two forms - binary and ascii.

A 1 Binary Format

This format was used for studying the characteristics of the GPS signals. The binary message has the following structure:

Field	Description of the Field	Size in bytes
'?' '?'	Indicates the beginning of the message	2
Message ID	Message identification	2
Body of the message	The actual GPS processed data	Variable length
Checksum	Checksum for all bytes upto the end of the above message	1
Line Feed	Indicates the end of message	1

1 Geodetic Position This message is used by the receiver to periodically transmit the position in the selected geodetic frame to the host. The transmitted message has the following format - ??AB001234CL

AB00 is the message id, C is the checksum noted above and L is the Linefeed. The message body has the following fields:

- 1 = Its a 4 bytes unsigned integer indicating the time tag in the GPS time format, i.e., the number of seconds since the beginning of the week (0 hrs UTC Sunday)
- 2 = 4 bytes Latitude in units of 10^{-7} degrees
- 3 = 4 bytes Longitude in units of 10^{-7} degrees
- 4 = 4 bytes Altitude in units of 0.01 meters

2 ECEF Position This is also a periodic message transmitted by the receiver giving the user's position but in terms of ECEF (Earth Centered Earth Fixed) coordinates The message has the following format - ??AC001234CL

AC00 is the message id while C and L as usual stand for the checksum and linefeed The details of the message body are

1 = A 4 byte time tag indicating the number of seconds since the beginning of the GPS week

2 = 4 bytes ECEF X coordinate in units of 0.01 meters

3 = 4 bytes ECEF Y coordinate in units of 0.01 meters

4 = 4 bytes ECEF Z coordinate in units of 0.01 meters

3 Dop Estimate This is a periodic transmit message which gives information about the satellite geometry of the satellites used for position solution In 2D mode only HDOP is transmitted and in 3D mode both HDOP and PDOP are transmitted The message has the following format - ??A0001234CL The details of the message body are as follows

1 = 2 bytes PDOP in units of 0.01

2 = 2 bytes HDOP in units of 0.01

3 = 2 bytes GDOP in units of 0.01

4 = 2 bytes error estimate in units of 0.1 meters

4 Error Message This is a periodic message which is transmitted by the receiver every 250 milliseconds and indicates any errors that have occurred in the receiver The message has the format - ??AA5512CL 1 and 2, each 2 bytes are unsigned integers indicating the type of errors occurred

A 2 Ascii Format

This format was used in setting up the time server The ascii message format that the receiver outputs conforms to the NMEA 0183 standard The NMEA (National Marine Electronics Association) 0183 belongs to the family of the Marine interface standards The NMEA 0183 standard specifies data communication in the form of coded sentences Each sentence begins with the character "\$" and ends with a

carriage return and linefeed (<CR> <LF>) Between the beginning and end of each sentence are “fields” of data, each field separated by a comma. The first field in any sentence begins with a two letter mnemonic code symbolizing the transmitter (Data emanating from a GPS receiver starts with “GP”) followed by a three letter code for that sentence. We used the \$GPRMC string output to get the required timing information. The format of the string is as follows

```
$GPRMC,hhmmss.ss,a ddmm mmmm,n,dddmm mmmm,w,zzz zz,yyy yy,ddmmyy
                                     *CC <CR><LF>
```

“hhmmss.ss” denote the hours, minutes and seconds (upto two decimals) respectively of the UTC time of position fix

“a” is a status indicator with “A” for valid and “V” for invalid

The next two fields give an estimate of the user’s latitude

dd	=	degrees (00 - 90)
mm mmmm	=	minutes (00 0000 - 59 9999)
n	=	direction (‘N’ for north and ‘S’ for south)

The next two fields provide the user longitude estimates

ddd	=	degrees (00 - 180)
mm mmmm	=	minutes (00 0000 - 59 9999)
w	=	direction (‘E’ for east and ‘W’ for west)

The next two fields give the user’s velocity and direction of heading

The next field gives the UTC date of position fix with ‘dd’ representing the date, ‘mm’ the month and ‘yy’ the two digit year

CC represents the checksum, calculated upto the above field

References

- [1] "NAVSTAR GPS Operations", [http // tycho usno navy mil/gps html](http://tycho.usno.navy.mil/gps.html)
- [2] "Global Positioning System Theory and Applications", volume I, edited by B W Parkinson and J J Spilker
- [3] "Understanding GPS Principles and Applications", Elliot D Kaplan Artech House
- [4] "GPSR 2000 Evaluation kit User's guide" supplied alongwith the Accords GPS Receiver
- [5] "Time Server" [http // www eecis udel edu/~ntp](http://www.eecis.udel.edu/~ntp)
- [6] J Steiner,B Neuman and J Schiller,"Kerberos An authentication service for open network systems", in proceedings of the usenix conference, Feburary 1988 (Available from [http // web mit edu/kerberos/www/papers html](http://web.mit.edu/kerberos/www/papers.html))
- [7] "Modelling and Analysis of computer network clocks", technical report, David L Mills, University of delaware (Available from [http // www eecis udel edu/~mills/ bib html](http://www.eecis.udel.edu/~mills/bib.html))
- [8] 'Network Time Protocol-version three specification implementation and analysis RFC 1305 David L Mills
- [9] David L Mills 'Internet time synchronization the Network Time Protocol ' IEEE Trans Communications 39, 10 (October 1991)
- [10] David L Mills Precision synchronization of computer network clocks" ACM computer Communication Review 24, 2 (April 1994)
- [11] David L Mills On the accuracy and stability of clocks synchronized by the Network Time Protocol in the Internet system", ACM Computer Communication Review 20, 1 (Januaury 1990)
- [12] "Network Time Protocol version four proposed changes", technical report David L Mills and A Thyagarajan, University of delaware
- [13] "Potential networking applications of Global positioning system", technical report, [http // www cis ohio-state edu/~jain/papers html](http://www.cis.ohio-state.edu/~jain/papers.html)

125737

A 125737
Date Slip

This book is to be returned on the
date last stamped

--

EE-1998-M-HEM-STU



A125737